





Query	Subject	Score	DB	Length	Ident	Gap	Gap
Q1	1 AGCTCTCAATACGGAACATATCTCAGTCTCTCTCTGCTCTACAAACCGTGTATCT	60.6%	DB 2	166873	100%	0	0
DB	26036 AGCTCTCAATATGGAACATATCTCAGTCTCTCTCTGCTCTACAAACCGTGTATCT	94.0%	Pred. No. 7.5e-58				
Matches	266; Conservative	0; Mismatches	15; Indels	2; Gaps	2;		
Q2	61 TGTCTATGACAGACGCTGTGTTAACTACAGGACCCAAACCTCTGTAAGCTTAT						
DB	26156 TGTCTATGACAGACGCTGTGTTAACTACAGGACCCAAACCTCTCTGTAAGCTTAT						
Q3	121 GCTTACAGTACACGTCGAGTCAGTCTGATTAATTTATTTGTTTCAGTCCTGAGATATA						
DB	26216 GCTTACAGTACACGTCGAGTCAGTCTGATTAATTTATTTGTTTCAGTCCTGAGATATA						
Q4	181 GCTACAGAACCTTTCCATTCGCATTCAGTACGAAACCTGCAAGGAGTGTGCTACAG						
DB	26276 GCTACAGAACCTTTCCATTCGCATTCAGTACGAAACCTGCAAGGAGTGTGCTACAG						
Q5	241 GTCTACAAAGAACTTCACATCATCTTCTTGAAGGAAAGAAC						
DB	26334 GTCTACAAAGAACTTCACATCATCTTCTTGAAGGAAAGAAC						
LOCUS	AC090848	169559 bp	DNA	linear	ORI	25,000	
DEFINITION	Homo sapiens chromosome 11, clone KPII-78024, complete sequence.						
VERSION	AC090848.4	GI:1795297					
KEYWORDS	HTG.						
SOURCE	Homo sapiens.						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.						
REFERENCE	1 (bases 1 to 169559)						
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.						
JOURNAL	Unpublished						
REFERENCE	2 (bases 1 to 169559)						
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,...						

[illegible][illegible]

repeat_L_region	complement(1350, 1535)
repeat_L_region	/rpt_family="L1HA10"
repeat_L_region	1713, 3831
repeat_L_region	/rpt_family="L1H4"
repeat_L_region	4030, 4311
repeat_L_region	/rpt_family="L12"
repeat_L_region	complement(4446, 5200)
repeat_L_region	/rpt_family="L1MC7"
repeat_L_region	complement(5201, 5472)
repeat_L_region	/rpt_family="A10S"
repeat_L_region	complement(5473, 5545)
repeat_L_region	/rpt_family="L1W7"
repeat_L_region	complement(5700, 5864)
repeat_L_region	/rpt_family="L1B7"
repeat_L_region	5894, 6295
repeat_L_region	/rpt_family="ML1B"
repeat_L_region	complement(6348, 6457)
repeat_L_region	/rpt_family="ML1A1"
repeat_L_region	7251, 7561
repeat_L_region	/rpt_family="L1MC"
repeat_L_region	7570, 7805
repeat_L_region	/rpt_family="L1MC7"
repeat_L_region	7862, 8200
repeat_L_region	/rpt_family="L1H4"
repeat_L_region	8202, 8766
repeat_L_region	/rpt_family="L1R22"
repeat_L_region	8951, 9677
repeat_L_region	/rpt_family="HERV822"
repeat_L_region	10047, 10215
repeat_L_region	/rpt_family="HERV822"
repeat_L_region	10621, 10656
repeat_L_region	/rpt_family="HERV822"
repeat_L_region	10696, 11735
repeat_L_region	/rpt_family="HERV822"
repeat_L_region	11805, 12768
repeat_L_region	/rpt_family="HERV822"
repeat_L_region	12770, 13336
repeat_L_region	/rpt_family="L1R22"
repeat_L_region	13371, 13494
repeat_L_region	/rpt_family="L1H4"
repeat_L_region	13567, 13565
repeat_L_region	/rpt_family="HERV51A"
repeat_L_region	14200, 14617
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	14820, 16078
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	16085, 18932
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	18937, 19341
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	19622, 19673
repeat_L_region	/rpt_family="HERV3"
repeat_L_region	19763, 20051
repeat_L_region	/rpt_family="L1R4"
repeat_L_region	20136, 20633
repeat_L_region	/rpt_family="HERV51A"
repeat_L_region	20813, 20878
repeat_L_region	/rpt_family="M1R"
repeat_L_region	complement(21660, 21798)
repeat_L_region	/rpt_family="HERV5A"
repeat_L_region	22492, 22591
repeat_L_region	/rpt_family="L1R35"
repeat_L_region	22714, 23046
repeat_L_region	/rpt_family="L1R35"
repeat_L_region	23456, 23816
repeat_L_region	/rpt_family="L2"
repeat_L_region	complement(23801, 23858)
repeat_L_region	/rpt_family="L2"
repeat_L_region	24476, 24582
repeat_L_region	/rpt_family="L2"
repeat_L_region	25186, 25278
repeat_L_region	/rpt_family="L1ME1"
repeat_L_region	25309, 26522

[illegible][illegible]





[illegible]





REFERENCE 1 (bases 1 to 1492)

Strausberg, K.

Submitted (05-NOV-2001) National Institutes of Health, Bethesda, MD, U.S.A.

USA  
Gene Collection (MGC), Cancer Genomics Unit, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD - 20892-0001

NIH-MGC Project URL: <http://mgc.bcl.nhl.gov>

CONTACT: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: AICU

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (1994)

RNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbiology.org>

contact: amaddan@systemsbiology.org

Anup Madan, Jessica Farley, Erin Helton, Mark Vetterlund, Elizabeth Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Witting

Clone distribution: MGC clone distribution information can be found through the J.M.A.G.E. Consortium/LiM at: <http://jmae.lim.gov>  
Series: IRAL Plate: 32 Row: j Column: 6

FEELINGS	Location/Qualifiers
Source	1. 1492

/orɣän ɪsm= "Homo sapiens"

```
/db_xref="taxon:9606"  
/clone="MGC:24011 IMAGE:"
```

```
/tissue_type="Bone marrow"
/clone_lib="NIH_MGC_55"
```

```

/lab_host="DH10B"
/note="Vector: pUNR-LIB"

```

```

;JS      276. .1160
         /codon_start=1

```

```

/product="Unknown (protein)"
/protein_id="AAH16785.1"

```

```
/db_xref="GI:16877032"  
/translation="MLVLHNSQKI"
```

LVDAWPDYQIVITRPQKQENKDDQ  
IOGCOGLDEAIRKVAITSKSVQV

GNESIMFDASHAGLVNEHWAFCB  
OSCELRMGYTPKRYRHQGNMLQIQ

КІСРСГННҚСТРКНУС"

[illegible]

Query Match	Score
30.0%;	Score 116.4
97.0%;	Read No. 1

Matches 128; Conservative 0; Mismatch

257 CAGATCATCTCTGAGGGAAGAACTGAAGTGG

Db 233 CAGATCATCTTCTGAGGGAAGAACCTGAGCTGG

317 C T C T C A G A G C T G C A G A T T C G T A T A A A T C C T T A G

De 253 CTCGAGAGCTGCAGATCTGTATAAATCCTTA

QY 377 ANJSTATNGCG 386

352 AUGTATATGCG 363  
 DD

REF ID: A64-315

LOC 2 AF429315 125020 bp  
DEFINITION Homo sapiens junctionin 3 (JPH

```

ACCESSION      AF429315
VERSION        AF429315.1  GI:17646244

```

KEYWORDS  
SOURCE  
Homo sapiens.

EUKARYOTISM    Homo sapiens  
Eukaryota; Metazoa; Chordata; Cr

QY 361 TCCCTGATCCATTAANGTAI 382  
DB 61021 DEGYSTWIDHMSKMBRD 61042

RESULT 11  
AP000576 121922 bp DNA Insert  
LOCUS Homo sapiens chromosome 11 clone CHB9-6J10 map 11q12 WORKING DRAFT  
DEFINITION SEQUENCE, 12 unordered pieces.  
ACCESSION AP000576.4 GI:9927276  
VERSION AP000576.4  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens DNA, clone:CHB9-6J10.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Catarrhini; Hominoidea; Homo.  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 121922)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,F.B., Hong,S.-G., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakai,Y.  
Homo sapiens 121,922 genomic DNA of 11q12  
Published Only in Database (1999)  
2 (bases 1 to 121922)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,F.B., Hong,S.-G., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakai,Y.  
Direct Submission  
Submitted (08-Oct-1999) Masahiro Hattori, The Institute of Chemical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-15-1 Kitasato, Sayamihara, Kanagawa 248-0555, Japan  
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgc.riken.go.jp/  
Tel:81-42-778-9923, Fax:81-42-778-9924)  
On Aug 26, 2000 this sequence version replaced gi:119274.  
----- Genome Center  
Center: RIKEN Genomic Sciences Center(GSC)  
Center code: RIKEN  
Web site: http://hgc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: HUMDRAFT1  
Center clone name: CHB9-6J10  
----- Summary Statistics  
Sequencing vector: PCR products; 1066 of reads  
Chemistry: Dye-terminator ET-amersham; 1008 of reads  
Assembly program: Phrap; version 0.990379  
Consensus quality: 118280 bases at least 40  
Consensus quality: 116659 bases at least 40  
Consensus quality: 120421 bases at least 40  
Insert size: 120822; sum-of-ctrls  
Quality coverage: 12.41x in 920 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N' but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 25319 contig of 25319 bp in length
25420 43536 contig of 18117 bp in length
43537 62163 contig of 18527 bp in length
62164 74470 contig of 12207 bp in length
74471 84826 contig of 10256 bp in length
84827 94411 contig of 9485 bp in length
94412 103560 contig of 9049 bp in length
103561 111843 contig of 8183 bp in length
111844 116835 contig of 4896 bp in length
116836 119647 contig of 2708 bp in length
119648 120811 contig of 1064 bp in length
120812 121922 contig of 1011 bp in length.

```

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces as is not known and their order in this sequence record is

```

FEATURES
Source      1..121922
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="11"
             /map="11q12"
             /clone="CHB9-6J10"
             1..25319
             /note="Assembly-Fragment"
             25420..44536
             /note="Assembly-Fragment"
             43537..62163
             /note="Assembly-Fragment"
             62164..74470
             /note="Assembly-Fragment clone_end:576 vector_side:left"
             74471..84826
             /note="Assembly-Fragment"
             84827..94411
             /note="Assembly-Fragment"
             94412..103560
             /note="Assembly-Fragment"
             103561..111843
             /note="Assembly-Fragment clone_end:11 vector_side:left"
             111844..116835
             /note="Assembly-Fragment"
             116836..119647
             /note="Assembly-Fragment"
             119648..120811
             /note="Assembly-Fragment"
             120812..121922
             /note="Assembly-Fragment"
BASE 1 3081 35022 A 21606 C 24325 G 37609 T 1190 others
ORIGIN
1 25319 contig of 25319 bp in length
25420 43536 contig of 18117 bp in length
43537 62163 contig of 18527 bp in length
62164 74470 contig of 12207 bp in length
74471 84826 contig of 10256 bp in length
84827 94411 contig of 9485 bp in length
94412 103560 contig of 9049 bp in length
103561 111843 contig of 8183 bp in length
111844 116835 contig of 4896 bp in length
116836 119647 contig of 2708 bp in length
119648 120811 contig of 1064 bp in length
120812 121922 contig of 1011 bp in length.

```

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. The true order of the pieces as is not known and their order in this sequence record is

```

RESULT 12
AP001636      161582 bp   DNA       linear    PF: 08-16-97
LOCUS         Homo sapiens genomic DNA, chromosome 11q, clone:RP11-142C4
DEFINITION   complete sequence.
ACCESSION    AP001636
VERSION      AP001636.4 GI:17425232
KEYWORDS     HTG.
SOURCE       Homo sapiens DNA, clone:RP11-142C4.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seock,P.,
              Fujiyama,A., Yada,T., Tokohi,Y., Watanabe,H. and Sakaki,Y.
TITLE        Homo sapiens genomic DNA
JOURNAL      Published Only in Database (2000)
REFERENCE    2 (bases 1 to 161582)
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seock,P.,
              Fujiyama,A., Yada,T., Tokohi,Y., Watanabe,H. and Sakaki,Y.
TITLE        Direct Submission
JOURNAL      Submitted (05-APR-2000) Masahiro Hattori, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 3680
              1-7-22 Tsuchino-Chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045 Japan
              (E-mail:hattori@gsc.riken.go.jp, tokohima_kanagawa@ipc.riken.go.jp /
              Tel:81-45-503-9111, Fax:81-45-503-9170)
              On Dec 7, 2001 this sequence version replaced gi:19927250.

COMMENT
FEATURES
             location/qualifiers
Source       1..161582
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="11"
              /map="11q"
              /clone="RP11-142C4"

BASE COUNT   48492 a 31636 c 32205 g 49245 t

ORIGIN
Query Match          11.6% Score 45.2; DB 9; Length 161582;
Best Local Similarity 68.8%; Pred. No. 0.022; Indels -; Gaps 1;
Matches 77; Conservative 0; Mismatches 33;

QY 255 TTTCAGTCACTCTTCTTGAGGAAGAAGAC--CGAAGTCGTACATAGACTGTTCTTTTC 31-
DB 46166 TTTCAGAGTTCTTCTTGTAATGGCTGCACAGACATGGAACCATCCACAATAATGATTCTATGA 4---3
QY 313 ATPACTCTAGAGAGCTGCAGATTCGTGATTAATCTCTTACAAAAGAGATACC 34-
DB 46226 ATAATGCCAGAGAGCTGCTGCTTCATACAAACCTTAGCTGGAGCATAC 45--7

RESULT 13
AC090309      161586 bp   DNA       linear    PF: 14-05-97
LOCUS         Homo sapiens chromosome 1l, clone RP11-142C4, complete sequence.
DEFINITION   complete sequence.
ACCESSION    AC090309
VERSION      AC090309.4 GI:15055316
KEYWORDS     HTG.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Landier,E., Allen,N., Anders N.S.,
              Barnes,N., Bastien,V., Boguslavsky,L., Boucknight B., Broxton A.,
              Camarata,J., Campopiano,A., Choquet,Y., Colangelo,M., Collins S.,
              Collumore,A., Cooke,P., DeArnellano,K., Dewar,K., Disteche,S.,
              Dodge,J.S., Fero,S., Ferreira,P., FitzHugh,W., Gage,F., Gallagher,

```

[illegible]

[illegible]







[illegible]

XX	Human: 1	State cancer: cytostatic; carcinogen; pharmacodynamic marker;
XX	plasma	marker; gene: SS.
CS	Homo # 4:	SS.
PD	W 2001	-78-
FD	23-Feb	-
XX		
PF	20-Feb	- 61W-BS05171.
XX		
PR	13-Feb	- 90HS-18313P.
FR	16-Mar	- 1 - 20HS-18962P.
FR	25-Mar	- 6 - 200S-207454P.
FR	09-Jun	- 6 - 2000S-211314P.
FR	18-Jul	- 6 - 2000S-215007P.
FR	13-Dec	- 6 - 2000S-235281P.
PA	(MILLI) : MEDENION PREDICTIVE MEDICINE INC.	
XX	Schreier,	Erlene E., Mosbach JE;
XX		
XX	WP# :	- 1027 of 775.
XX		
P1	Novel 19	nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer; useful for detecting presence of prostate cancer; stage of prostate cancer -
FI		
FI	Claim 1	stage 1027: 117504P; English.
PS		
CC	The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in tables 1-9 (ABV0010-ABV62213) or the specification or its complement. (1) is useful for:	
CC	(a) assessing whether a patient is afflicted with prostate cancer;	
CC	(b) monitoring the progression of prostate cancer in a patient;	
CC	(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;	
CC	(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;	
CC	(e) determining a composition for inhibiting prostate cancer in a patient;	
CC	(f) assessing the prostate cell carcinogenic potential of a compound;	
CC	(g) detecting whether prostate cancer has metastasized in a patient;	
CC	(h) assessing the aggressiveness or incidence of prostate cancer in a patient;	
CC	(1) is also useful as a pharmacodynamic or pharmacogenomic marker.	
XX		
SV	Sequence: 1190 bp; 384 A; 328 C; 350 G; 321 T; 7 other:	
	Query Match:	111st Score 43; DB 23; Length 1900;
	Best local similarity:	58.6%; Pval: 0.0042;
	Matches:	% Conservativity of Mismatches 32; Indels 1; gaps 1
CY	423 CTGACAGCAATGGTCTGTGTCATCATCCTCGAAGACTGCACATCTGTATA 342	
EL	352 GTTTTAAATGAATGATGCTGTGTAATGATGCCATACACCTCTCTGCTATATA 421	
WJ	343 AAAATGAAAATGATGCTGTGTAATGATGAAAATGATATTAC 367	
LQ	422 AAATGATGCAAGAGATACCTGATGACCTGATGAGAGTGATATGCG 465	
RESULT 5		
ABV21557		
ID	ABV21557	(standard); CDNA: 1390 BP.
XX	ABV21557	
XX	1c-SEF - 2 (first entry)	
EE	Human 1:	tate expression marker CDNA 21548.
XX	Human 1:	state cancer; cytostatic; carcinogen; pharmacodynamic marker;



AM pharmacogenomic marker; gene; ss.  
 XX  
 US Homo sapiens.  
 XX  
 PH WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 FE 20-FEB-2001: 2001WO-US05171.  
 XX  
 FR 17-FEB-2000: 2000US-183319P.  
 XX  
 FR 16-MAR-2000: 2000US-189862P.  
 XX  
 FR 25-MAY-2000: 2000US-207454P.  
 XX  
 FR 09-JUN-2000: 2000US-211314P.  
 XX  
 FR 18-JUL-2000: 2000US-215007P.  
 XX  
 FR 13-DEC-2000: 2000US-255281P.  
 XX  
 PH (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 LR WPI: 2001-662795/76.  
 XX  
 PI Novel isolated nucleic acid molecule associated with cancerous state of  
 PI prostate cells and correlating with presence of prostate cancer, useful  
 PI for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1: Page 4654; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (i) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (i) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 XX  
 SN Sequence 1390 BP: 384 A; 328 C; 350 G; 321 T; 7 other:  
 XX  
 Query Match 11.1% Score 43: DB 23: Length 1390:  
 Best Local Similarity 68.6% Pred. No. 0.0042:  
 Matches 72: Conservative 0: Mismatches 32: Indels 1: Gaps 1:  
 XX  
 CY 283 CTGAGTGCATCAATGAGTCTGCTGCTCAATGCTCTGAGAGAGCTGAGATTCGTATTA 342  
 XX  
 PD 362 CTGAGATCCACAGAGATGATCTACTGAACTCAATGAGCTGCTGCTGATATA 421  
 XX  
 CY 343 AATCTTAGAAGAGAGATCCCTGGAATCCATAAGATATATTCG 387  
 XX  
 PD 422 AATCTTGCCAGAGAGCAT-CCCTGAGTCCCTGAAGCGTATATGC 465  
 XX  
 RESULT 6  
 ABV24715  
 ID ABV24715 standard; cDNA: 1390 BP.  
 XX  
 AC ABV24715:  
 XX  
 AT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 24706.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX

S Homo sapiens.  
 XX  
 PH WO 200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 FE 20-FEB-2001: 2001WO-US05171.  
 XX  
 FR 17-FEB-2000: 2000US-183319P.  
 XX  
 FR 16-MAR-2000: 2000US-189862P.  
 XX  
 FR 25-MAY-2000: 2000US-207454P.  
 XX  
 FR 09-JUN-2000: 2000US-211314P.  
 XX  
 FR 18-JUL-2000: 2000US-215007P.  
 XX  
 FR 13-DEC-2000: 2000US-255281P.  
 XX  
 PH (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege W, Monahan JE;  
 XX  
 LR WPI: 2001-662795/76.  
 XX  
 PI Novel isolated nucleic acid molecule associated with cancerous state of  
 PI prostate cells and correlating with presence of prostate cancer, useful  
 PI for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1: Page 4713; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (i) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (i) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 XX  
 SN Sequence 1390 BP: 384 A; 328 C; 350 G; 321 T; 7 other:  
 XX  
 Query Match 11.1% Score 43: DB 23: Length 1390:  
 Best Local Similarity 68.6% Pred. No. 0.0042:  
 Matches 72: Conservative 0: Mismatches 32: Indels 1: Gaps 1:  
 XX  
 CY 283 CTGAGTGCATCAATGAGTCTGCTGCTCAATGCTCTGAGAGAGCTGAGATTCGTATTA 342  
 XX  
 PD 362 CTGAGATCCACAGAGATGATCTACTGAACTCAATGAGCTGCTGCTGATATA 421  
 XX  
 CY 343 AATCTTAGAAGAGAGATCCCTGGAATCCATAAGATATATTCG 387  
 XX  
 PD 422 AATCTTGCCAGAGAGCAT-CCCTGAGTCCCTGAAGCGTATATGC 465  
 XX  
 RESULT 6  
 ABV25372  
 ID ABV25372 standard; cDNA: 1390 BP.  
 XX  
 AC ABV25372:  
 XX  
 AT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 25372.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX



XX	21-MAR-2002.
XX	10-SEP-2001: 2001WO-US26015.
XX	11-SEP-2000: 2000US-0659671.
XX	(HYSE-) HYSEQ INC.
XX	Tang YF, Liu C, Zhou P, Asundi V, Zhang J, Zhao CA Ref: 1
P1	Xue AJ, Yang Y, Wehrman I, Demanac RT:
DR	WP1: 2002-292408/33.
DR	P-PSDB: ABB57238.
P1	An isolated polynucleotide for treating diseases associated with its
FI	encoded polypeptide such as cancer and multiple sclerosis -
XX	Claim 1: SEQ ID NO 62; 505pp; English.
XX	The present invention provides the protein and coding sequences of 44;
CC	novel human proteins. These were isolated from expressed sequences lars
CC	(ESIs). They can be used to stimulate cell growth, to regulate
CC	haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC	e.g. in burn treatment, to regulate the immune system e.g. to treat
CC	multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC	fertility, to regulate haemostasis or thrombolysis e.g. to treat
CC	stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC	e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC	Parkinson's disease. The present sequence is a coding sequence of the
CC	invention.
SQ	Sequence 1627 BP; 472 A; 336 C; 394 G; 420 T; 5 other:
XX	Query Match 11.1%; Score 43; DB 24; Length 1627;
XX	Best Local Similarity 68.6%; Pred. No. 0.0044;
XX	Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps
C:	263 CTGAGTCTCATTAAGATGCTTGCTTTAACTCTCAGAGCTTCAGATTCTGATAA
DE	310 CTGAAGCATCCACAGATGTACTCTACTGATATACCTCCCAATCTCTGGTATACA
C:	343 AATCCTTAGAAAAGACATCCCCTGAAATGCATAAATGATAATMC 367
DE	370 AATCCTTGCCAGSAGCAT-CCCTGAGTCCTGAAGGTGATATGCC 413
RESULT 10	
AAH23819	
ID	AAH23819 standard; cDNA; 1955 bp.
AC	AAH23819;
DT	31-JUL-2001 (first entry)
XX	Human transferase HTFS-19 cDNA, SEQ ID NO:61.
XX	Human transferase: HTFS; agonist; antagonist; cellular stimulation;
KM	proliferation; cell proliferative disorder; immune disorder;
KM	atherosclerosis; hepatitis; psoriasis; cancer; tumour;
KM	inflammation; AIDS; Addison's disease; allergy; asthma; arthritis;
KM	cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
KM	multiple sclerosis; rheumatoid arthritis; pancreatitis;
KM	systemic lupus erythematosus; thymocytopenia; ulcerative colitis;
KM	haemodialysis; extracorporeal circulation; trauma; transgenic animal;
XX	gene therapy; drug screening; ss.
XX	Homo sapiens.
OS	
FH	Key Location/Qualifiers
FT	CDS 270..1178
FT	/tag-a
FT	/product= "HTFS-19"
FT	/function= "transferase"

[illegible]

```

XX ABV42132:
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker CDNA 42123.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker:
XX pharmacogenomic marker; gene: ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001: 2001WO-US05171.
XX
XX 17-FEB-2000: 2000US-183319P.
XX 16-MAR-2000: 2000US-185862P.
XX 25-MAY-2000: 2000US-207454P.
XX 05-JUN-2000: 2000US-211314P.
XX 18-JUL-2000: 2000US-219007P.
XX 13-DEC-2000: 2000US-255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI: 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1: Page 8445; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (i) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV21213) of the
XX specification or its complement. (1) is useful for:
XX
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 358 BP: 90 A; 85 G; 92 G; 97 T; 0 other:
XX
XX Query Match 10.7%: Score 41.4; DB 23; Length 358;
XX Best Local Similarity 67.6%: Pred. No. 0.0077;
XX Matches 71: Conservative 0; Mismatches 33; Indels 1; gaps
XX
XX 263 CTGAAGTGCATACATAGATGCTTGCTTCATACACTCTCGAGACGCGAGATTCTGATATA 344
XX 11111 11 111 1 11 11111 1 11111 1 1111 1 1111 1
XX Db 122 CTGAGACGATCCCAAGATATGATGCTACTGATATACATCCCATCAACTGCTGGAGCTATATA 18;
XX
XX 343 AATCC11AGAAAGACGATCCCTGATATCATTAAGAATATATAC 367
XX 1111111 1 1111 11 11111111 1 11 11 11
XX Db 182 AATCTTGCGCCAGGAGTAT-CCTGTAGTCCCTGAGGTGATATGCG 225
XX
XX RESULT 12
XX ID ABV22508 standard; cDNA; 368 BP.
XX
XX ABV22908:
XX

```

[illegible]

```

XX DE Human prostate expression marker CDNA 22903.
XX KM Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker.
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN W0200160860-A2.
XX FD 23-AUG-2001.
XX PE 20-FEB-2001; 2001W0-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-185862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-215007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE.
XX DR WPI: 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1: Page 4060; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) determining whether prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
XX CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SO Sequence 388 BP; 94 A; 77 C; 100 G; 113 T; 4 other:
XX
XX Query Match 10.4%; Score 40.4; DB 23; Length 368;
XX Best Local Similarity 66.1%; Pred. No. 0.016;
XX Matches 74; Conservative 0; Mismatches 36; Indels 2; Gaps 1;
XX
XX QY 255 TTGAGTCACTCTCTGAGGAAGAAG--CTGAGTGTACTATAAGATCTGTGCTG 31-
XX DB 93 TTGAGATTCTCTCTCAAGGTCTCAAGGTCTCAAGATCCACAGATATATCTCTG 15-
XX QY 313 ATAACTCTGAGAGTGCACATCTGTATAATCTTAAAGAGGATATCC 364
XX DB 153 ATAACTCCATAGTGTGCGCCCTATACAAATCTCTTGGCAGAGGATCC 204
XX
XX RESULT 14
XX ID ABV28742 standard; cDNA; 388 BP.
XX AC ABV28742;
XX XX 16-SEP-2002 (first entry)
XX DE Human prostate expression marker CDNA 28733.

```

```

XX DE Human prostate expression marker CDNA 13719.
XX KM Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker.
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN W0200160860-A2.
XX FD 23-AUG-2001.
XX PE 20-FEB-2001; 2001W0-US05171.
XX PR 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-185862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-215007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE.
XX DR WPI: 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1: Page 6037; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (1) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (1) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) determining whether prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient;
XX CC (1) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SO Sequence 388 BP; 94 A; 77 C; 100 G; 113 T; 4 other:
XX
XX Query Match 10.4%; Score 40.4; DB 23; Length 368;
XX Best Local Similarity 66.1%; Pred. No. 0.016;
XX Matches 74; Conservative 0; Mismatches 36; Indels 2; Gaps 1;
XX
XX QY 255 TTGAGTCACTCTCTGAGGAAGAAG--CTGAGTGTACTATAAGATCTGTGCTG 31-
XX DB 93 TTGAGATTCTCTCTCAAGGTCTCAAGGTCTCAAGATCCACAGATATATCTCTG 15-
XX QY 313 ATAACTCTGAGAGTGCACATCTGTATAATCTTAAAGAGGATATCC 364
XX DB 153 ATAACTCCATAGTGTGCGCCCTATACAAATCTCTTGGCAGAGGATCC 204
XX
XX RESULT 15
XX ID ABV13728 standard; cDNA; 382 BP.
XX AC ABV13728;
XX XX 16-SEP-2002 (first entry)
XX DE Human prostate expression marker CDNA 13719.
XX KM Human: prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

```

pharmacogenomic marker; gene; ss.

homo sapiens.

W0200160860-A2.

23-AUG-2001.

20-FEB-2001: 2001WO-US05171.

17-FEB-2000: 2000US-183319P.

16-MAR-2000: 2000US-189862P.

25-MAY-2000: 2000US-207454P.

09-JUN-2000: 2000US-211314P.

18-JUL-2000: 2000US-219007P.

13-DEC-2000: 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI: 2001-662795/76.

Claim 1: Page 2279; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62313) or the specification or its complement. (I) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient;

(d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient;

(e) selecting a composition for inhibiting prostate cancer in a patient;

(f) assessing the prostate cell carcinogenic potential of a compound;

(g) determining whether prostate cancer has metastasized in a patient;

(h) assessing the aggressiveness or indolence of prostate cancer in a patient;

(i) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 392 BP: 96 A; 77 C; 102 G; 114 T; 3 other;

Query Match 10.4%; Score 40.4; DB 23; Length 392;

Best Local Similarity 66.1%; Pred. No. 0.016; Mismatches 36; Indels 2; Gaps 1;

Matches 74; Conservative 0; Mismatches 36; Indels 2; Gaps 1;

255 TTGAGATCATCTTCTTGGAGGAAGAG--CTGAGTGTCTAGATAGATCTTGTGCTTC 312

97 TTGAGATCTTCTTCTTGGAGGAAGAG--CTGAGTGTCTAGATAGATCTTGTGCTTC 156

313 ATAGCTCTAGAGATCTTCTTGGAGGAAGAG--CTGAGTGTCTAGATAGATCTTGTGCTTC 354

157 ATAGCTCTAGAGATCTTCTTGGAGGAAGAG--CTGAGTGTCTAGATAGATCTTGTGCTTC 208

Search completed: January 2, 2003, 13:48:57

Job time : 232 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

nm nucleic - nucleic search, using sw model

run on: January 2, 2003, 13:11:27 : Search time 1945 seconds  
(without alignments)  
3224.140 Million cell updates/sec

Title: US-09-762-027-4  
Perfect score: 388  
Sequence: 1 agccgcacatagcaaacat.....atccataatgatalatgcy j88

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 809774376 residues  
Total number of hits satisfying chosen parameters: 32408132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estopl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	235	60.6	786	17	AQ750247 HS-5576_A
2	138.8	35.8	720	12	BF214297 60184619
3	105.4	27.2	814	12	BF240961 60187551
4	47.6	12.3	709	12	BG430720 60249161
5	45.4	11.7	914	12	BG430555 602502206
6	43	11.1	376	9	AA340632 EST15952

Result No.	Score	Query Match	Length	DB ID	Description
1	502	100	502	1	AF159347
2	498	100	498	1	AF159347
3	494	100	494	1	AF159347
4	490	100	490	1	AF159347
5	486	100	486	1	AF159347
6	482	100	482	1	AF159347
7	478	100	478	1	AF159347
8	474	100	474	1	AF159347
9	470	100	470	1	AF159347
10	466	100	466	1	AF159347
11	462	100	462	1	AF159347
12	458	100	458	1	AF159347
13	454	100	454	1	AF159347
14	450	100	450	1	AF159347
15	446	100	446	1	AF159347
16	442	100	442	1	AF159347
17	438	100	438	1	AF159347
18	434	100	434	1	AF159347
19	430	100	430	1	AF159347
20	426	100	426	1	AF159347
21	422	100	422	1	AF159347
22	418	100	418	1	AF159347
23	414	100	414	1	AF159347
24	410	100	410	1	AF159347
25	406	100	406	1	AF159347
26	402	100	402	1	AF159347
27	398	100	398	1	AF159347
28	394	100	394	1	AF159347
29	390	100	390	1	AF159347
30	386	100	386	1	AF159347
31	382	100	382	1	AF159347
32	378	100	378	1	AF159347
33	374	100	374	1	AF159347
34	370	100	370	1	AF159347
35	366	100	366	1	AF159347
36	362	100	362	1	AF159347
37	358	100	358	1	AF159347
38	354	100	354	1	AF159347
39	350	100	350	1	AF159347
40	346	100	346	1	AF159347
41	342	100	342	1	AF159347
42	338	100	338	1	AF159347
43	334	100	334	1	AF159347
44	330	100	330	1	AF159347
45	326	100	326	1	AF159347

## ALIGNMENTS

Result 1  
Query: 576-A2-E06-SPE-KF1-11 Human Male BAC library; Homo sapiens  
Definition: 1152 Col-10 Row=1, DNA sequence.

Accession: AF159347  
Version: 1  
Keywords: NUCLEIC ACID, REPEATS, REPLICATION ORIGIN, REPLICATION

## REFERENCE

1. Miller, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Feldman, R., Shaver, R., Furlong, J., Young, J., Zhao, S., Adams, M.L. and  
J.L. (1999) A sequence approach to mapping and  
identifying the human genome  
2. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
3. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
4. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
5. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
6. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
7. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
8. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
9. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
10. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
11. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
12. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
13. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
14. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
15. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
16. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
17. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
18. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
19. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
20. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
21. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
22. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
23. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
24. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
25. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
26. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
27. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
28. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
29. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
30. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
31. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
32. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
33. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
34. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
35. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
36. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
37. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
38. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
39. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
40. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
41. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
42. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
43. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
44. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
45. J. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)





source

1. .814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4091916"  
 /clone\_lib="NIH\_MGC\_55"  
 /tissue\_type="from acute myelogenous leukemia"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: bone marrow; Vector: pBluescript (Clontech);  
 site\_1: SfiI (ggcgctcgccg); site\_2: SfiI (ggcgctcgccg)  
 5' and 3' adaptor cDNA was prepared from cell line RNA.  
 5' and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CAGCGCCATATGCGC-3' and 3'  
 adaptor sequence: 5'-ATCTAGAGCGCCGAGCGCGGATG-3' (B30)BN-3'  
 (where B = A, C, G, or T and N = A, C, G, or T). Average  
 insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA)."

BASE COUNT 255 a 175 c 191 g 169 t

ORIGIN

Query Match 27.28; Score 105.4; DB 12; Length 911;  
 Best Local Similarity 96.28; Pred. No. 1.1e-20;  
 Matches 128; Conservative 0; Mismatches 3; Indels 2; Gaps 4;

QY 257 CAGATGATCTCTTGAGGAGGAGAGAGCTAGCTAGTACATAATGCT-TGCTTTCATTA 315  
 B3430730  
 LOCUS 6023493681F1 NIH\_MGC\_75 Homo sapiens cDNA clone IMAGE:4607516 5'  
 DEFINITION mRNA sequence.

ACCESSION B3430730  
 VERSION B3430730  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 709)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Clontech Laboratories, Inc.  
 CDNA Library Preparation: Clontech Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHCN1348 row: a column: 21  
 High quality sequence stop: 514.  
 Location/Qualifiers  
 1..709  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4607516"  
 /clone\_lib="NIH\_MGC\_75"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /lab\_host="NIH\_MGC\_75"  
 /note="Organ: kidney; Vector: pDR-LIB (Clontech); Site\_1:

site\_2: SfiI (ggcgctcgccg); site\_2: SfiI (ggcgctcgccg); 5' and  
 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CAGCGCCATATGCGC-3' and 3'  
 adaptor sequence: 5'-ATCTAGAGCGCCGAGCGCGGATG-3' (B30)BN-3'  
 (where B = A, C, G, or T and N = A, C, G, or T). Average  
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained  
 inserts by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: This is a NIH-MGC library."

BASE COUNT 206 a 145 c 157 g 199 t

ORIGIN

Query Match 12.93; Score 47.6; DB 12; Length 763;  
 Best Local Similarity 77.05; Pred. No. 0.003;  
 Matches 62; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 350 AAGCTTGGCTTCTGATGAGGAGGAGAGAGCTAGCTAGTACATAATGCT-TGCTTTCATTA 359  
 B3430555  
 LOCUS 602502206F1 NIH\_MGC\_75 Homo sapiens cDNA clone IMAGE:4515625 5'  
 DEFINITION mRNA sequence.

ACCESSION B3430555  
 VERSION B3430555  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 714)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
 Tissue Procurement: Clontech Laboratories, Inc.  
 CDNA Library Preparation: Clontech Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LHCN1365 row: 1 column: 02  
 High quality sequence stop: 645.  
 Location/Qualifiers  
 1..714  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4515625"  
 /clone\_lib="NIH\_MGC\_75"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: kidney; Vector: pDR-LIB (Clontech); site\_1:  
 SfiI (ggcgctcgccg); site\_2: SfiI (ggcgctcgccg); 5' and  
 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CAGCGCCATATGCGC-3' and 3'  
 adaptor sequence: 5'-ATCTAGAGCGCCGAGCGCGGATG-3' (B30)BN-3'  
 (where B = A, C, G, or T and N = A, C, G, or T). Average  
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained  
 inserts by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: This is a NIH-MGC library."

BASE COUNT 236 a 200 c 210 g 239 t

ORIGIN

Query Match 10.03; Score 45.4; DB 12; Length 911;







AUTHORS NIH-MGC <http://imgc.ncl.nih.gov/>  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://imgc.llnl.gov>  
 Plate: LLCM1330 row: f column: 09  
 High quality sequence stop: 695.  
 Location/Qualifiers

FEATURES  
 source  
 1..703  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4592264"  
 /clone\_1lb="NIH-MGC\_75"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /note="Organ: Kidney; Vector: pMR-1lb (Clontech); Site\_1: SfiI (ggcgcttcggcc); Site\_2: SfiI (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCAATATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGCCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT  
 origin 203 a 151 c 168 g 181 t

Query Match 11.18; Score 43; DB 12; Length 703;  
 Best Local Similarity 68.6%; Pred. No. 0.074;  
 Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

263 CTGAGTGTCTCATAGTCTGCTGCTCACTCACTCAAGCTCAGATTCGATA 342  
 241 CTGAGGATCCACAGATGATCTACTGATATACCTCCCAAGCTGTGCCCTATACA 300

343 AATCCTTAGAAGAGCATCCCTGGAATCCATAAAGTATATATGCC 367  
 301 AATCCTTGCCAGAGGAGCAT-CCCTGAGTCCCTGAGAGGTATATGCC 344

RESULT 14  
 B3400739 734 bp mRNA linear EST 12-MAR-2001  
 LOCUS 602461127F1 NIH-MGC\_75 Homo sapiens cDNA clone IMAGE:4592264  
 DEFINITION mRNA sequence.  
 ACCESSION B3400739  
 VERSION B3400739  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC <http://imgc.ncl.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://imgc.llnl.gov>  
 Plate: LLCM1330 row: e column: 08

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

High quality sequence stop: 726.  
 Location/Qualifiers  
 1..734  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4592264"  
 /clone\_1lb="NIH-MGC\_75"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /note="Organ: Kidney; Vector: pMR-1lb (Clontech); Site\_1: SfiI (ggcgcttcggcc); Site\_2: SfiI (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCAATATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGCCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT  
 origin 210 a 153 c 170 g 185 t

Query Match 11.18; Score 43; DB 12; Length 734;  
 Best Local Similarity 68.6%; Pred. No. 0.076;  
 Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

263 CTGAGTGTCTCATAGTCTGCTGCTCACTCACTCAAGCTCAGATTCGATA 342  
 241 CTGAGGATCCACAGATGATCTACTGATATACCTCCCAAGCTGTGCCCTATACA 300  
 343 AATCCTTAGAAGAGCATCCCTGGAATCCATAAAGTATATATGCC 367  
 301 AATCCTTGCCAGAGGAGCAT-CCCTGAGTCCCTGAGAGGTATATGCC 344

RESULT 15  
 B3428541 744 bp mRNA linear EST 14-MAR-2001  
 LOCUS 602501064F1 NIH-MGC\_75 Homo sapiens cDNA clone IMAGE:4614687  
 DEFINITION mRNA sequence.  
 ACCESSION B3428541  
 VERSION B3428541  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC <http://imgc.ncl.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://imgc.llnl.gov>  
 Plate: LLCM1330 row: f column: 16  
 High quality sequence stop: 706.  
 Location/Qualifiers

FEATURES  
 source  
 1..744  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4614687"  
 /clone\_1lb="NIH-MGC\_75"  
 /lab\_host="DH10B (11 phage-resistant)"  
 /note="Organ: Kidney; Vector: pMR-1lb (Clontech); Site\_1: SfiI (ggcgcttcggcc); Site\_2: SfiI (ggcattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCCAATATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGCCGCCGACATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT 203 a 159 c 169 g 212 t 1 others  
ORIGIN

Query Match 11.18; Score 43; DB 12; Length 744;  
Best Local Similarity 68.68; Pred. NO. 0.076;

Matches 72; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 283 CTGAGTCTACATPAGATCTGCTTCATPACTCTCAGAGCTCCAGATTCCTGATA 342

Db 382 CTGAGCATCCACAGATGATCTCTACTGAATTAATCCCATAGCTGCTGCGCCATACA 441

QY 343 AATCCTTAGAAAGAGCATCCCTGATCCATTAANGTATATTCG 387

Db 442 AATCCTTGCCAGGAGCAT-CCCTGAGTCCCTGAGAGGTGATGGC 485

Search completed: January 2, 2003, 14:21:43  
Job time: 1955 secs

GenCore version 5.1.3  
Copyright (c) 1995 - 2003 CompuGen Ltd.

us-09-762-027-4

Run on: January 2, 2003, 13:44:02 : Search time 47 seconds  
(without alignments)  
2531.715 Million cell updates/sec

Title: US-09-762-027-4  
Perfect score: 388  
Sequence: 1 agcgcctcaatcagcaacat.....atccataatgataatgacg jbb

Scoring table: IDENTIFY NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338481 residues

Total number of hits satisfying chosen parameters: 843724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCUTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfillseq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32.2	8.3	307	4	US-09-172-108-5
2	31	8.0	1249	2	US-08-933-750C-62
3	31	8.0	1249	3	US-05-231-613-62
4	31	8.0	1802	3	US-05-289-377-1
5	30.8	7.9	1578	4	US-09-134-001C-677
6	30.8	7.9	5181	4	US-08-257-073-10
7	30.2	7.8	16950	4	US-05-453-702B-160
8	30	7.7	1661	1	US-08-518-474-1
9	29.6	7.6	295	4	US-09-172-711-52
10	29.4	7.6	758	2	US-08-720-258-1
11	29.4	7.6	2119	2	US-08-381-691-17
12	28.8	7.4	2174	4	US-09-367-206-4
13	28.2	7.3	500	3	US-09-141-000-2
14	28	7.2	512	2	US-08-967-101-36
15	28	7.2	512	2	US-08-592-541-36
16	28	7.2	512	3	US-09-124-698-36
17	28	7.2	512	4	US-09-127-480-36
18	28	7.2	512	4	US-08-496-841C-36
19	28	7.2	512	4	US-09-124-523-36
20	28	7.2	699	4	US-09-134-001C-1609
21	28	7.2	5761	4	US-09-323-472A-1
22	28	7.2	5761	4	US-09-323-472A-3
23	28	7.2	5761	4	US-09-323-472A-11
24	28	7.2	5762	4	US-09-323-472A-13
25	27.8	7.2	552	2	US-08-614-377A-12
26	27.8	7.2	552	4	US-05-142-648B-12
27	27.8	7.2	777	2	US-08-720-258-3

1	Sequence 1, Appl
2	Sequence 7, Appl
3	Sequence 3, Appl
4	Sequence 1, Appl
5	Sequence 3, Appl
6	Sequence 35, Appl
7	Sequence 35, Appl
8	Sequence 35, Appl
9	Sequence 35, Appl
10	Sequence 3, Appl
11	Sequence 10, Appl
12	Sequence 1315, Ap
13	Sequence 2815, Ap
14	Sequence 8, Appl
15	Sequence 195, App
16	Sequence 18, Appl
17	Sequence 170, App
18	Sequence 1, Appl

## ALIGNMENTS

1	Sequence 1, Appl
2	Sequence 7, Appl
3	Sequence 3, Appl
4	Sequence 1, Appl
5	Sequence 3, Appl
6	Sequence 35, Appl
7	Sequence 35, Appl
8	Sequence 35, Appl
9	Sequence 35, Appl
10	Sequence 3, Appl
11	Sequence 10, Appl
12	Sequence 1315, Ap
13	Sequence 2815, Ap
14	Sequence 8, Appl
15	Sequence 195, App
16	Sequence 18, Appl
17	Sequence 170, App
18	Sequence 1, Appl

Fri Jan 3 14:27:02 2003

us-09-762-027-4.rni

Page 2

APPLICANT: Yue, Henry  
 APPLICANT: Guejler, Karl J.  
 APPLICANT: Corley, Neil C.  
 TITLE OF INVENTION: HUMAN REGULATORY MOLECULES  
 NUMBER OF SEQUENCES: 98  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: diskette  
 COMPUTER: IBM compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSB0 for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/933,750C  
 FILING DATE: September 23, 1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0356 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555  
 TELEFAX: 415-845-4166  
 TELEX:  
 INFORMATION FOR SEQ. ID NO.: 62:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1249 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: SPINNOT02  
 CLONE: 208636  
 US-06-933-750C-62

```

Query Match      8.0% Score 31: DB 2: Length 125%:
Best Local Similarity 57.3%: Pred. No. 2:
Matches 53: Conservative 0: Mismatches 41: Indels 0: Gaps
CY      258  GTGCACATAGATGCTGCTTCATCAATATTCAGAAAGTCAGATTCCTGATTAATATCC 347
      1 11111 111 1111111 11 11111 1111 1111 1111 1
Db      1122  GAGCAGACGAGAGAGGCTGGCTCTCCAGCTTGGCCAGAAAGACCCGATGAGCATGCAATTC 112
      1 11111 111 1111111 111 1111 1111 1111 1111 1
CY      348  TTAGAAAGAGCATCCCTGATTCATTAANGTATA 383
      1 1 11 111 111 111 111 111 111 111 111 111
Db      1182  CCATTACTTAGGTTTCCCATATTCCTTAATTAATAA 1217
      1 1 11 111 111 111 111 111 111 111 111 111

RESULT 3
US-09-234-613-62
: Sequence 62, Application US/09234613
: Patent No. 6132973
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Purvi
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

```

1 STREET: 1114 PATER: 1100
2 CITY: PALM BEACH
3 STATE: FLA
4 COUNTRY: USA
5 ZIP: 33404
6
7 COMPUTER MAKE/ABLE: IBM:
8
9 MODEL: 4741: 4741: 4741:
10
11 OPERATING SYSTEM: DOS
12
13 SOFTWARE: EASYSEEK VOF Windows Version 2.0
14
15 CURRENT APPLICATION DATA:
16
17 APPLICATION NUMBER: 05/02/14 013
18
19 FILED DATA:
20
21 CLASSIFICATION:
22
23 PRIOR APPLICATION DATA:
24
25 APPLICATION NUMBER: 05/02/13 750
26
27 FILING DATE: September 23, 1997
28
29 ATTORNEY/AGENCY INFORMATION:
30
31 NAME: BILIMORE, LUCY J.
32
33 REGISTRATION NUMBER: 35,719
34
35 REFERENCE/DOCKET NUMBER: PR-055 US
36
37 TELECOMMUNICATION INFORMATION:
38
39 TELEPHONE: 415-855-0555
40
41 TELEFAX: 415-843-4145
42
43 TELE:
44
45 INFORMATION FOR SERIAL NO. 14:
46
47 SEQUENCE CHARACTERISTICS:
48
49 LENGTH: 149 base pairs
50
51 TYPE: nucleic acid
52
53 STRANDEDNESS: single
54
55 TOPOLGY: linear
56
57 IMMEDIATE SOURCE:
58
59 LIBRARY: SPIN102
60
61 GENE: 200036
62
63 US-05-234-613-02

```

```

Query Match Similarity 57.3%: Fred No. 2;
Header Local Similarity 57.3%: Fred No. 2;
Matches 55: Conservative 0: Mismatches 41: Indels 0: Gaps 0

266 GAGTACACAAATACCTATCTCTATACCTCAGAACTCGAGCTCTGTATATAC 347
1 11111 111 11111 1 1111 111 111 1
1112 GAGTACACAAATACCTATCTCTATACCTCAGAACTCGAGCTCTGTATATAC 1161
1 11111 111 11111 1 1111 111 111 1

348 TATATAAAAGATACCTCGATACCAATATATATAT 383
1 11111 111 111 111 111 11 1
1162 GAGTACACAAATACCTATCTCTATACCTCAGAACTCGAGCTCTGTATATAC 1217
1 11111 111 11111 1 1111 111 111 1

Result 4
US-09-289-377-1/3
Sequence 1 Application US-09-289-377
Patent No. 0940351
GENERAL INFORMATION:
APPLICANT: C. M. CORSE
FILE OF INVENTION: ANTISENSE REGULATION OF G-ALPHA-11 EXPRESSION
FILE REFERENCE: R15-0055
CURRENT APPLICATION NUMBER: US-09/289, 377
CURRENT FILING DATE: 1999-04-09
NUMBER OF SEQ. ID NOS: 47
SEQ. ID NO. 1
LENGTH: 1802
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (13)...(177)
US-09-289-377-1

6.0%: Score 31; DB 3; Length 1202;
Best Match Similarity 62.0%: Pref. No. 2, 3;
Matches 49: Conservative 0: Mismatches 30: Indels 0: Gaps 0

```







```

ADDRESS: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple System 7.5.3
SOFTWARE: Microsoft Word, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,258
FILING DATE: 26-SEPT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/575,715
FILING DATE: 20-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,324
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2620-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756832
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Compox p35
FEATURE:
NAME/KEY: CDS
LOCATION: 18..758
NAME/KEY: mat_peptide
LOCATION: 66..755
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 18..65
US-08-720-258-1
Query Match 7.6%; Score 29.4; Db 2; Length 758;
Best Local Similarity 58.6%; Pred. No. 5.1;
Matches 51; Conservative 0; Mismatches 36; Indels 0; Gaps 0
QY 216 CTGGACGACGACGCTGCTACACGCTCAGACAGAAATTCAGATCATCTTTGAGGG 275
DB 48 CTGGGACGACGCTGCTACACGCTCAGACAGATCATCTTTGAGGG 275
QY 276 AAGAGCTGAAGTGTCTACATAGATG 302
DB 108 ACGGAGAGAGAAACAAACATCATATG 134
RESULT 11
US-08-381-691-17
Sequence 17, Application US/08381691
Patent No. 5852224
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Alpha-lac Albumin Gene Constructs
NUMBER OF SEQUENCES: 17
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Patchin, release 3.0, Version 31.30 (lib)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/381,691
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2119 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-381-691-17
Query Match 7.6%; Score 29.4; Db 2; Length 2119;
Best Local Similarity 51.1%; Pred. No. 7.9;
Matches 09; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 105 CTCTGGAGCTTTTATCTTATACAGTGCACAGTGCATTCATTTATTTGTTTTC 164
DB 25 CTCTGGAGCTCTCTGAGTGCATTCATTCATTCATTCATTCATTCATTCATTCATTC 85
QY 105 AGTGGTAAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 224
DB 09 TCAATATATATCTCTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 145
QY 146 GGTATATGCTATCA 239
DB 146 TATATATCTCTCA 190
RESULT 12
US-09-367-206-1
Sequence 1, Application US/09367206
Patent No. 625162
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: HSP molecules
FILE REFERENCE: P1223RIE
CURRENT APPLICATION NUMBER: US/09/367,206
FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/08847
PRIOR FILING DATE: 1999-04-23
PRIOR APPLICATION NUMBER: US 60/082,767
PRIOR FILING DATE: 1996-04-23
PRIOR APPLICATION NUMBER: US 60/113,296
PRIOR FILING DATE: 1996-12-24
NUMBER OF SEQ ID NOS: 35
SEQ ID NO: 1
LENGTH: 2171
TYPE: DNA
ORGANISM: Homo sapiens
US-09-367-206-1
Query Match 7.4%; Score 28.8; Db 4; Length 2171;
Best Local Similarity 56.2%; Pred. No. 12;
Matches 51; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 421 AGATATAGAGTGGTCTGAGTTCATTCATTCATTCATTCATTCATTCATTCATTC 280
DB 104 AGATATAGAGTGGTCTGAGTTCATTCATTCATTCATTCATTCATTCATTCATTC 1763
QY 421 AGATATAGAGTGGTCTGAGTTCATTCATTCATTCATTCATTCATTCATTCATTC 316
DB 1764 AGATATAGTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1799
RESULT 13
US-09-141-000-270
Sequence 2, Application US/09141000

```



Fri Jan 3 14:27:02 2003

us-09-762-027-4.rni

Page 7

QY 336 CTGTATAATCCTAGA 352  
|||||  
Db 353 CTGTCTNTATGCTNAGA 369

Search completed: January 2, 2003, 15:06:17  
Job time : 56 secs





QY 280 AAGCTGAGTGTACATAGATGCTGTCTTCACTACATCTCAGAGCTGAGATTTCGT 335  
 DB 58 AAGCTTACGTTGCTGATGATGATGCTTCCCTCCAGAGTGCAGATGCTGAGATTTCGT 157  
 QY 340 ATAAATCCTTAGAAAAGACATCC 364  
 DB 158 AGAAATCCTTAGAGAGAGACCTTC 182

# RESULT 6

US-09-960-352-7792  
 : Sequence 7792, Application US/09960352  
 : Patent No. US20020137139A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Warren, Wesley C.  
 : APPLICANT: Tao, Nengbing  
 : APPLICANT: Byatt, John C.  
 : APPLICANT: Mathialagan, Nagappan  
 : TITLE OF INVENTION: MUSCLE ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 : FILE REFERENCE: 16511.006/37-21(10298)C  
 : CURRENT APPLICATION NUMBER: US/09/960.352  
 : CURRENT FILING DATE: 2001-05-24  
 : NUMBER OF SEQ ID NOS: 15112  
 : SEQ ID NO: 7792  
 : LENGTH: 313  
 : TYPE: DNA  
 : ORGANISM: bos taurus  
 : OTHER INFORMATION: Clone ID: 33-L1B34-076-Q1-E1-A2  
 US-09-960-352-7792

Query Match  
 Best Local Similarity 64.7%; Pred. No. 0.044;  
 Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0

QY 280 AAGCTGAGTGTACATAGATGCTGTCTTCACTACATCTCAGAGCTGAGATTTCGT 335  
 DB 58 AAGCTTACGTTGCTGATGATGATGCTTCCCTCCAGAGTGCAGATGCTGAGATTTCGT 157  
 QY 340 ATAAATCCTTAGAAAAGACATCC 364  
 DB 158 AGAAATCCTTAGAGAGAGACCTTC 182

# RESULT 7

US-09-960-352-14120  
 : Sequence 14120, Application US/09960352  
 : Patent No. US20020137139A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Warren, Wesley C.  
 : APPLICANT: Tao, Nengbing  
 : APPLICANT: Byatt, John C.  
 : APPLICANT: Mathialagan, Nagappan  
 : TITLE OF INVENTION: MUSCLE ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 : FILE REFERENCE: 16511.006/37-21(10298)C  
 : CURRENT APPLICATION NUMBER: US/09/960.352  
 : CURRENT FILING DATE: 2001-05-24  
 : NUMBER OF SEQ ID NOS: 15112  
 : SEQ ID NO: 14120  
 : LENGTH: 303  
 : TYPE: DNA  
 : ORGANISM: bos taurus  
 : OTHER INFORMATION: Clone ID: 60-L1B34-027-Q1-E1-C8  
 US-09-960-352-14120

Query Match  
 Best Local Similarity 64.7%; Pred. No. 0.048;  
 Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0

QY 280 AAGCTGAGTGTACATAGATGCTGTCTTCACTACATCTCAGAGCTGAGATTTCGT 335  
 DB 58 AAGCTTACGTTGCTGATGATGATGCTTCCCTCCAGAGTGCAGATGCTGAGATTTCGT 157

QY 340 ATAAATCCTTAGAAAAGACATCC 364  
 DB 212 AATAATCCTTAGAGAGAGACCTTC 240

# RESULT 8

US-09-960-352-7790  
 : Sequence 7790, Application US/09960352  
 : Patent No. US20020137139A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Warren, Wesley C.  
 : APPLICANT: Tao, Nengbing  
 : APPLICANT: Byatt, John C.  
 : APPLICANT: Mathialagan, Nagappan  
 : TITLE OF INVENTION: MUSCLE ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 : FILE REFERENCE: 16511.006/37-21(10298)C  
 : CURRENT APPLICATION NUMBER: US/09/960.352  
 : CURRENT FILING DATE: 2001-05-24  
 : NUMBER OF SEQ ID NOS: 15112  
 : SEQ ID NO: 7790  
 : LENGTH: 313  
 : TYPE: DNA  
 : ORGANISM: bos taurus  
 : OTHER INFORMATION: Clone ID: 29-L1B34-011-Q1-E1-B1  
 US-09-960-352-7790

Query Match  
 Best Local Similarity 64.7%; Pred. No. 0.049;  
 Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0

QY 280 AAGCTGAGTGTACATAGATGCTGTCTTCACTACATCTCAGAGCTGAGATTTCGT 335  
 DB 183 AAGCTTACGTTGCTGATGATGATGCTTCCCTCCAGAGTGCAGATGCTGAGATTTCGT 242  
 QY 340 ATAAATCCTTAGAAAAGACATCC 364  
 DB 243 AATAATCCTTAGAGAGAGACCTTC 267

# RESULT 9

US-09-960-352-7770  
 : Sequence 7770, Application US/09960352  
 : Patent No. US20020137139A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Warren, Wesley C.  
 : APPLICANT: Tao, Nengbing  
 : APPLICANT: Byatt, John C.  
 : APPLICANT: Mathialagan, Nagappan  
 : TITLE OF INVENTION: MUSCLE ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
 : FILE REFERENCE: 16511.006/37-21(10298)C  
 : CURRENT APPLICATION NUMBER: US/09/960.352  
 : CURRENT FILING DATE: 2001-05-24  
 : NUMBER OF SEQ ID NOS: 15112  
 : SEQ ID NO: 7770  
 : LENGTH: 303  
 : TYPE: DNA  
 : ORGANISM: bos taurus  
 : OTHER INFORMATION: Clone ID: 29-L1B34-001-Q1-E1-B1  
 US-09-960-352-7770

Query Match  
 Best Local Similarity 64.7%; Pred. No. 0.049;  
 Matches 55; Conservative 0; Mismatches 30; Indels 0; Gaps 0

QY 280 AAGCTGAGTGTACATAGATGCTGTCTTCACTACATCTCAGAGCTGAGATTTCGT 335  
 DB 183 AAGCTTACGTTGCTGATGATGATGCTTCCCTCCAGAGTGCAGATGCTGAGATTTCGT 242





QY 340 ATAAATCCTTAGAAGACCATCCC 364  
 | | | | | | | | | | | | | | | | | |  
 Db 223 AGAATCCTTGATGAGAGACCTTCC 247

## RESULT 14

US-09-770-445-209

Sequence 209: Application US/09770445

Patent No. US200202281A1

GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn

APPLICANT: An, Yong-Qiang

APPLICANT: Hamilton, Carol M.

APPLICANT: Price, Jennifer L.

APPLICANT: Raines, Tracy M.

APPLICANT: Yu, Yang

APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy

APPLICANT: Mathew, Abraham V.

APPLICANT: Ledford, Brooke L.

APPLICANT: Moesner, Jeffrey P.

APPLICANT: Haas, William David

APPLICANT: Garcia, Carlos A.

APPLICANT: Kricke, Neja

APPLICANT: Slader, Ted

APPLICANT: Davis, Keith R.

APPLICANT: Allen, Keith

APPLICANT: Hoffman, Neil

APPLICANT: Hurban, Patrick

APPLICANT: Title of Invention: Expressed Sequences of Arabidopsis

FILE REFERENCE: 2023US (PARA-012PRV)

CURRENT APPLICATION NUMBER: US/09770,445

CURRENT FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: US 60/178,472

PRIOR FILING DATE: 2000-01-27

NUMBER OF SEQ ID NOS: 999

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 209

LENGTH: 1022

TYPE: DNA

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(1022)

OTHER INFORMATION: n = A,T,C or G

US-09-770-445-209

Query Match

Best Local Similarity 8.5%; Score 32.8; DB 10; Length 1022;

Matches 76; Conservativity 0; Mismatches 72; Indels 0; Gaps 0;

QY 213 AAATGACAGACAGCTGCTACAGCTCAAGAACTTCATCTTCCTGA 272

| | | | | | | | | | | | | | | | | |

Db 12 AAGCTGTGAAGAGATTATATACACCAAAAAGACATGATTCAGCTTGG 71

QY 273 GGAAGAAGCTGAAGTGTACATAGATGCTGTCTTATACCTCAGAGCTTCAG 332

| | | | | | | | | | | | | | | | | |

Db 72 AAGACAGAGCTGAAGTGTATTAATTAAGATTTTGTACGAAGTTCCTACCAAAACT 131

QY 333 ATTCTGATTAATCTCTAGAAAGAGCA 360

| | | | | | | | | | | | | | | | | |

Db 132 TTATTATCAAAATGAAAAACAAAGAA 159

## RESULT 15

US-09-796-692-5227/C

Sequence 5227: Application US/09796692

Publication No. US20020198362A1

GENERAL INFORMATION:

APPLICANT: Gaigier, Alexander

APPLICANT: Gaigier, Paul A.

APPLICANT: Mannion, Jane

TITLE OF INVENTION: NEW STRAINS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

FILE REFERENCE: 2007-00-000000

CURRENT APPLICATION NUMBER: US/09/796,692

CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/185,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/200,545

PRIOR FILING DATE: 2000-03-27

PRIOR APPLICATION NUMBER: 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: 60/200,599

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: 60/218,950

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: 60/222,903

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: 60/229,416

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: 60/223,376

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 9797

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 527

LENGTH: 533

TYPE: DNA

ORGANISM: H. sapiens

US-09-796-692-5227

Query Match

Best Local Similarity 5.3%; Score 32.2; DB 9; Length 533;

Matches 70; Conservativity 0; Mismatches 63; Indels 0; Gaps 0;

QY 32 TCTCTTCTGTAAGAGTGTATCTGCTATGACAGACAGCTGCTTATCTA 91

| | | | | | | | | | | | | | | | | |

Db 339 TCCCTTCTGTAAGAGTGTATCTGCTATGACAGACAGCTGCTTATCTA 280

QY 92 CAGGATCTGTAAGAGTGTATCTGCTATGACAGACAGCTGCTTATCTA 151

| | | | | | | | | | | | | | | | | |

Db 279 AAGCTCTGTAAGAGTGTATCTGCTATGACAGACAGCTGCTTATCTA 220

QY 152 TTATTATCAAAATGAAAAACAAAGAA 159

| | | | | | | | | | | | | | | | | |

Db 219 TTATTATCAAAATGAAAAACAAAGAA 159

Search completed: January 4, 2003, 15:07:16  
 Job time: 53 s



```

1d 157 lysstlser-----AsPTATrYgSerSettlethuclyvallenleasplentylet 174
Cy 156 TCCATTGCC-----ATACTGAGAAACTGGCAGACAGGACGTGCTTCAAGTTTC 245
Cd 175 AAlaGAlaATrPGluIleValArgAsnValThcInIleLeuleu-----191
Cy 247 AAAGAAGCTTCAGATCATCTTGAGCGAAGAACCGTAGAGTGTfAcwIAAATGTTG 305
Cd 192 ---GlutserGlylleSerPheleuLusAnlysylsien---AlatrylleNeAsple 209
Cy 307 TGC 309
Cd 210 Cys 210

RESULT 2
Ab1505
Probable TNF2-like transposon protein [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: AB4505
Refin. X.: Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, U.D.; Follis, G.A.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Sheen, J.; Vankken, S.E.; Chayman, E.; Talbot, D.; Weiss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A. Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: AB4505
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1040 <STO>
A:Cross-references: GB:ABE02093; NID:g4432794; PIDN:AAD20645.1; USDB:Chou139
C:Genetics:
A:Gene: Atg413000
A:Map position: 2

Alignment Scores:
Fred. No.: 3.56 Length: 1040
Score: 68.00 Matches: 19
Percent Similarity: 41.258 Conservative: 14
Best Local Similarity: 23.758 Mismatches: 29
Query Match: 10.358 Indels: 16
DB: Caps: 3

US-09-762-v27-4 (1-388) x AB4505 (1-1040)

Cy 103 TCAGTCGGTAGATATTAGCTACAGAAACCTTTCATTGCCATGAGAAAGTGGAA 442
Cd 350 SerMetMetThrIleLysAlaAspAsnAsnIleuSerGluLysCysMetAspSerTrpAla 359
Cy 223 CAGGCATG-----TGCCTAAAGAT 243
Cd 370 GluIleuIleLysGluTrpLeuProProAspAsnIleSerAlaLuseryPrpGluIle 389
Cy 244 TACAAGAAAGCTTCAGATCATCTTCTTGAGCGAAGAAAGCTGAG--TCTGAAGAAA 305
Cd 390 GluLysLeuValSerSerHisGlyLeuProSerGluMetLeuAspValCysIleAsnTy 402
Cy 301 TGCTTGAGCTTC-----ATAACTCTCAGAGAGCTGAGCAATTCGTAATAATGCT 348
Cd 410 CysMetIlePheTrpGlyAspAspValAsnLeuGluIleLucysAtyPheCysGlyTrp 422

RESULT 3
A44850
E6 protein - human papillomavirus type 66
C:Species: human papillomavirus type 66
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 22-May-1993
C:Accession: A44850
Refin.X.: Beaudenon, S.; Favre, M.; Orth, G.
J. Clin. Microbiol. 29, 2656-2660, 1991
A>Title: Characterization of human papillomavirus type 66 from an invasive carcinoma of the cervix.
A:Reference number: A44850; MUID:92129556; PMID:1663515
A:Accession: A44850

```

[illegible]

```

QY      22  TITCATCTCTCCCTCGTCAGTATACAAAGCCTGATTTCTGT-----AAGGG    72
DB      39  TyrosylasertlierperserglylleProIlellelaalaalaamMcaspHrileidy 58
        ::::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      70  ACAGAAGCTGCTGTTTAATCTACAGGAACCAACTCTCTG-----AAGG    117
DB      59  ThrPheGlutMetValLysSerLeuSerLysPheAsnIleleThrAlavalHisLysTy 78
        ||| :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     118  IATGCTTAACGATGCACAACGTGACAGTGATGAAATTATATGGTGTTCACAGCGACAGTA 177
DB      73  TyrsrPheGlu-----GluTrpLysAsnIleValCysLeuSerSerLysGlnIle 26
        ||| ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     178  TTACGTACAGAACACTTTCACATGCC-----                204
DB      56  LeuasnHisValIleValSerIleGlyThrSerAsnIleaspPheLeuLysIleLysLys 117
        ||| ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     205  -----ATACGTAGACAACCTGCAGCACGCCAGCTGCTGCTCCATCACAGTATAAAGAACTTCA 253
DB     116  IlePheLeuleuSerSerGuluLeuLysTrpIleLysIleaspValIleasnGluTrpSer 135
        ::||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     255  GATCATCTCTTCTTAGGCGAAGAAGCTG 285
DB     136  GluHisIleValSerPheLeuLysLeu 144
        ::||| ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
T06019 .
hypothetical protein T25K17.160 - Arabidopsis thaliana
Cispeptides: Arabidopsis thaliana (mouse-ear cress)
CDate: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Nov-1999
CAccession: T06019
R.Subvan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, L.; Mexes, H.W.,
submitted to the Protein Sequence Database, March 1999
A.Reference number: Z15382
A.Accession: T06019
A.Molecule type: DNA
A.Residues: 1-431 <BEV>
A.Cross-references: EMBL:AL049171; GenBank:U00062; ATSP:T25K17.160
A.Experimental source: cultivar Columbia; BAC clone T25K17
C.Gene: ATSP:T25K17.160
A.Map position: 4
A.Introns: 289/3; 334/3

Alignment Scores:
Pred. No.:          7.15           Length:       431
Score:              65.50          Matches:       20
Percent Similarity: 46.32%         Conservative: 16
Best Local Similarity: 27.37%      Mismatches:   32
Query Match:        9.97%          Indels:       19
DB:                 2             Gaps:         6

US-05-762-027-4 (1-388) x T06019 (1-431)
QY      67  TCGACAGAA---CGTCGTGTTATCTACAGAACCAACTACTCTCGAAG-----CTT 117
DB      257  TrpValIgluaIalagIlelaValIthGlyIySerHisLysTrpLeuLysGluIleTr 276
        ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     118  TATGCTTAACGATGCACAACGTGACAGTGAAATTATATGGTGTTCACGTCTACAGTA 177
DB      277  TyrAla-----LysGlyLeuArgValCysLeuSerValSerGluVal 250
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     178  TTACGTACAGAACCTTTCATTCGATATGCATCTGTGAAACGCGGACAGGACAGTCTGTA 237
DB      291  MetAsnProTryAspMetCilePheHisIleLeuValaspLeuThrIleLysIleCysThr 310
        ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     238  CAGGCTTACAAAGAAACTTCAGATCTATCTCTTCAAGGAAG--AACTGAAGTGTAC 294
DB      311  GInglyTrpTrpAspLeuLeuThrHisMetLeuGInglySerProLysLeuArg----- 328
        ||| ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     295  ATAAGATGCTTGCTTCATATACCTTCAGAAC-----TGCANA 333
DB      329  -----PheIleuThrLeuThrAsnAspHisLysAsp 338
        ||| ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

[illegible]

```

A:Accession: S27008
A:Molecule type: DNA
A:Residues: 1-138 'C',140-269 <GRA3>
A:Cross-references: EMBL:X65014; NID:q3563; PIDD:CA44614e.1; FID:q3563
A:Experimental source: strain 07173
R:Grack, H.R.; Grohmann, L.; Choll, T.
FEBS Lett. 242, 4-8, 1988
A>Title: Mitochondrial ribosomes of yeast: isolation of individual proteins and determination of their amino acid sequence
A:Accession: S26756
A:Molecule type: Protein
A:Residues: 21, V, 22-53 <GRA>
A:Experimental source: strain 07173
A:Note: This sequence has been revised in reference S23455
R:Grohmann, L.; Grack, H.R.; Kruffy, V.; Choll, T.; Goldschmidt-Beisler, S.; Kitzawa, M
FEBS Lett. 264, 51-56, 1991
A>Title: Extended N-terminal sequencing of proteins of the large ribosomal subunit from yeast
A:Reference number: S17255; MUID:89076818; PMID:3060376
A:Accession: S26757
A:Molecule type: Protein
A:Residues: 176, YP, 179-182, 257-261 <GRD>
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
submitted to the EMBL Data Library, June 1995
A:Description: Sequence analysis of the 43 KB CERN-YLM9-FE154-SM11-PH.31-YM94-FE154-100
A:Reference number: S57680
A:Accession: S57683
A:Molecule type: DNA
A:Residues: 1-269 <VAN>
A:Cross-references: EMBL:X87941; NID:q886508; PIDD:CA611e.1; FID:q886508
A:Experimental source: strain S286C
R:van der Aart, Q.J.M.; Steensma, H.Y.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64541
A:Accession: S64544
A:Molecule type: DNA
A:Residues: 1-269 <VAN>
A:Cross-references: EMBL:J73004; NID:q1323594; PIDD:CAAY724d.1; FID:q1323594; QEPID:CAAY724d.1
A:Experimental source: strain S286C
R:van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.
test 12, 385-390, 1996
A>Title: Sequence analysis of the 43 KB CERN-YLM9-FE154-DHE2-SM11-PH.31-YM94-FE154-100
A:Reference number: S63896; MUID:96267763; PMID:8701610
A:Accession: S63899
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-269 <VAN>
A:Cross-references: EMBL:X87941; NID:q886508; PIDD:CA611e.1; FID:q886508
A:Note: the nucleotide sequence was submitted to the EMBL Data Library June 1995
A:Genes: SCD:MRPL9; YML5; MIPS:YGR220C
A:Cross-references: SCD:S0003452; MIPS:YGR220C
A:Map position: 7R
A:Genome: nuclear
A:Superfamily: Escherichia coli ribosomal protein L3
A:Keywords: mitochondrion; protein biosynthesis; ribosome
F1-19/Domains: transit peptide (mitochondrion) *status experimental *status experimental *status experimental *status experimental
F1-20-269/Product: ribosomal protein L3, mitochondrial *status experimental *status experimental *status experimental *status experimental

```

[illegible]

1. Characterization of the Cephalosporin acetylase from *Streptomyces clavuligerus*.  
 2. Site of early cephalosporin biosynthetic genes and evidence of multiple horizontal gene  
 3. Transfer number: A38531; MUID:51177827; PMID:1706706  
 4. Accession: A38531  
 5. Molecule type: DNA  
 6. Accession: J01177  
 7. Accession: J01177  
 8. Accession: J01177  
 9. Accession: J01177  
 10. Accession: J01177  
 11. Accession: J01177  
 12. Accession: J01177  
 13. Accession: J01177  
 14. Accession: J01177  
 15. Accession: J01177  
 16. Accession: J01177  
 17. Accession: J01177  
 18. Accession: J01177  
 19. Accession: J01177  
 20. Accession: J01177  
 21. Accession: J01177  
 22. Accession: J01177  
 23. Accession: J01177  
 24. Accession: J01177  
 25. Accession: J01177  
 26. Accession: J01177  
 27. Accession: J01177  
 28. Accession: J01177  
 29. Accession: J01177  
 30. Accession: J01177  
 31. Accession: J01177  
 32. Accession: J01177  
 33. Accession: J01177  
 34. Accession: J01177  
 35. Accession: J01177  
 36. Accession: J01177  
 37. Accession: J01177  
 38. Accession: J01177  
 39. Accession: J01177  
 40. Accession: J01177  
 41. Accession: J01177  
 42. Accession: J01177  
 43. Accession: J01177  
 44. Accession: J01177  
 45. Accession: J01177  
 46. Accession: J01177  
 47. Accession: J01177  
 48. Accession: J01177  
 49. Accession: J01177  
 50. Accession: J01177  
 51. Accession: J01177  
 52. Accession: J01177  
 53. Accession: J01177  
 54. Accession: J01177  
 55. Accession: J01177  
 56. Accession: J01177  
 57. Accession: J01177  
 58. Accession: J01177  
 59. Accession: J01177  
 60. Accession: J01177  
 61. Accession: J01177  
 62. Accession: J01177  
 63. Accession: J01177  
 64. Accession: J01177  
 65. Accession: J01177  
 66. Accession: J01177  
 67. Accession: J01177  
 68. Accession: J01177  
 69. Accession: J01177  
 70. Accession: J01177  
 71. Accession: J01177  
 72. Accession: J01177  
 73. Accession: J01177  
 74. Accession: J01177  
 75. Accession: J01177  
 76. Accession: J01177  
 77. Accession: J01177  
 78. Accession: J01177  
 79. Accession: J01177  
 80. Accession: J01177  
 81. Accession: J01177  
 82. Accession: J01177  
 83. Accession: J01177  
 84. Accession: J01177  
 85. Accession: J01177  
 86. Accession: J01177  
 87. Accession: J01177  
 88. Accession: J01177  
 89. Accession: J01177  
 90. Accession: J01177  
 91. Accession: J01177  
 92. Accession: J01177  
 93. Accession: J01177  
 94. Accession: J01177  
 95. Accession: J01177  
 96. Accession: J01177  
 97. Accession: J01177  
 98. Accession: J01177  
 99. Accession: J01177  
 100. Accession: J01177  
 101. Accession: J01177  
 102. Accession: J01177  
 103. Accession: J01177  
 104. Accession: J01177  
 105. Accession: J01177  
 106. Accession: J01177  
 107. Accession: J01177  
 108. Accession: J01177  
 109. Accession: J01177  
 110. Accession: J01177  
 111. Accession: J01177  
 112. Accession: J01177  
 113. Accession: J01177  
 114. Accession: J01177  
 115. Accession: J01177  
 116. Accession: J01177  
 117. Accession: J01177  
 118. Accession: J01177  
 119. Accession: J01177  
 120. Accession: J01177  
 121. Accession: J01177  
 122. Accession: J01177  
 123. Accession: J01177  
 124. Accession: J01177  
 125. Accession: J01177  
 126. Accession: J01177  
 127. Accession: J01177  
 128. Accession: J01177  
 129. Accession: J01177  
 130. Accession: J01177  
 131. Accession: J01177  
 132. Accession: J01177  
 133. Accession: J01177  
 134. Accession: J01177  
 135. Accession: J01177  
 136. Accession: J01177  
 137. Accession: J01177  
 138. Accession: J01177  
 139. Accession: J01177  
 140. Accession: J01177  
 141. Accession: J01177  
 142. Accession: J01177  
 143. Accession: J01177  
 144. Accession: J01177  
 145. Accession: J01177  
 146. Accession: J01177  
 147. Accession: J01177  
 148. Accession: J01177  
 149. Accession: J01177  
 150. Accession: J01177  
 151. Accession: J01177  
 152. Accession: J01177  
 153. Accession: J01177  
 154. Accession: J01177  
 155. Accession: J01177  
 156. Accession: J01177  
 157. Accession: J01177  
 158. Accession: J01177  
 159. Accession: J01177  
 160. Accession: J01177  
 161. Accession: J01177  
 162. Accession: J01177  
 163. Accession: J01177  
 164. Accession: J01177  
 165. Accession: J01177  
 166. Accession: J01177  
 167. Accession: J01177  
 168. Accession: J01177  
 169. Accession: J01177  
 170. Accession: J01177  
 171. Accession: J01177  
 172. Accession: J01177  
 173. Accession: J01177  
 174. Accession: J01177  
 175. Accession: J01177  
 176. Accession: J01177  
 177. Accession: J01177  
 178. Accession: J01177  
 179. Accession: J01177  
 180. Accession: J01177  
 181. Accession: J01177  
 182. Accession: J01177  
 183. Accession: J01177  
 184. Accession: J01177  
 185. Accession: J01177  
 186. Accession: J01177  
 187. Accession: J01177  
 188. Accession: J01177  
 189. Accession: J01177  
 190. Accession: J01177  
 191. Accession: J01177  
 192. Accession: J01177  
 193. Accession: J01177  
 194. Accession: J01177  
 195. Accession: J01177  
 196. Accession: J01177  
 197. Accession: J01177  
 198. Accession: J01177  
 199. Accession: J01177  
 200. Accession: J01177  
 201. Accession: J01177  
 202. Accession: J01177  
 203. Accession: J01177  
 204. Accession: J01177  
 205. Accession: J01177  
 206. Accession: J01177  
 207. Accession: J01177  
 208. Accession: J01177  
 209. Accession: J01177  
 210. Accession: J01177  
 211. Accession: J01177  
 212. Accession: J01177  
 213. Accession: J01177  
 214. Accession: J01177  
 215. Accession: J01177  
 216. Accession: J01177  
 217. Accession: J01177  
 218. Accession: J01177  
 219. Accession: J01177  
 220. Accession: J01177  
 221. Accession: J01177  
 222. Accession: J01177  
 223. Accession: J01177  
 224. Accession: J01177  
 225. Accession: J01177  
 226. Accession: J01177  
 227. Accession: J01177  
 228. Accession: J01177  
 229. Accession: J01177  
 230. Accession: J01177  
 231. Accession: J01177

[illegible]







Genotex version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 2, 2003, 14:01:48 : Search time 15 seconds  
(without alignments)  
2475.818 Million cell updates/sec

File: US-09-762-027-4  
Perfect score: 657  
Sequence: 1 agctgcatacagacat.....atccataaanglulstnagc jcc

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 245744

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_model -DEV=xlh  
-C=/C:/usr2/spool/US09762027/runct\_02012003\_085613\_2543/app-query.fasta:553  
-DB=SwissProt\_40 -OPMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -L=0.01 -V  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -HITS=human40.tbl  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=20000000  
-USER=US09762027 -SCN\_1\_1\_6 -runat\_02012003\_085613\_2542 -NC=0 -IC=0.3  
-NO\_XMAP -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGJOB -EV=116204-1  
-NARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=0 -FGAPEXT=7  
-XNAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: SwissProt\_40:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.5	10.3	807	1	Y991_SCHPO
2	66	10.0	155	1	VE6_HPV66
3	66	10.0	349	1	GUAC_BUCAT
4	64.5	9.8	269	1	RM03_YEAST
5	64	9.8	215	1	U520_CAEEL
6	63.5	9.7	371	1	ACV5_CEPAC
7	63.5	9.7	120	1	SV16_HUMAN
8	62.5	9.5	471	1	CSN1_HUMAN
9	62.5	9.5	1400	1	R1P1_SCHPO
10	62	9.4	437	1	Y269_HELPY
11	62	9.4	437	1	Y269_HELPY
12	61.5	9.4	192	1	GCH2_HELPY
13	61	9.3	286	1	VNS2_INCYA
14	61	9.3	455	1	YHFN_ECOLI
15	61	9.3	599	1	HM21_HUMAN
16	60.5	9.2	1564	1	N184_SCHPO
17	60.5	9.2	164	1	CGHB_CALJA
18	60.5	9.2	252	1	Y060_METJA

12	60.5	9.2	252	1	Y060_METJA
13	61	9.3	286	1	VNS2_INCYA
14	61	9.3	455	1	YHFN_ECOLI
15	61	9.3	599	1	HM21_HUMAN
16	60.5	9.2	1564	1	N184_SCHPO
17	60.5	9.2	164	1	CGHB_CALJA
18	60.5	9.2	252	1	Y060_METJA
19	60.5	9.2	252	1	Y060_METJA
20	60.5	9.2	252	1	Y060_METJA
21	60.5	9.2	252	1	Y060_METJA
22	60.5	9.2	252	1	Y060_METJA
23	60.5	9.2	252	1	Y060_METJA
24	60.5	9.2	252	1	Y060_METJA
25	60.5	9.2	252	1	Y060_METJA
26	60.5	9.2	252	1	Y060_METJA
27	60.5	9.2	252	1	Y060_METJA
28	60.5	9.2	252	1	Y060_METJA
29	60.5	9.2	252	1	Y060_METJA
30	60.5	9.2	252	1	Y060_METJA
31	60.5	9.2	252	1	Y060_METJA
32	60.5	9.2	252	1	Y060_METJA
33	60.5	9.2	252	1	Y060_METJA
34	60.5	9.2	252	1	Y060_METJA
35	60.5	9.2	252	1	Y060_METJA
36	60.5	9.2	252	1	Y060_METJA
37	60.5	9.2	252	1	Y060_METJA
38	60.5	9.2	252	1	Y060_METJA
39	60.5	9.2	252	1	Y060_METJA
40	60.5	9.2	252	1	Y060_METJA
41	60.5	9.2	252	1	Y060_METJA
42	60.5	9.2	252	1	Y060_METJA
43	60.5	9.2	252	1	Y060_METJA
44	60.5	9.2	252	1	Y060_METJA
45	60.5	9.2	252	1	Y060_METJA

## ALIGNMENTS

## RESULT 1

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA

Y991\_SCHPO 2,400,000 607 AA



```

EMBL: AP001118; BAB12921.1;
HSSP: P49058; ILEP.
InterPro: IPR003009; ERM_enzyme.
InterPro: IPR001093; IMPDH/GMPase.
Pfam: PF00478; IMPDH_C.1.
TIGRfams: TIGR01305; GMP_reduct_1.1.
PROSITE: PS00487; IMP_DH_GMP_RED.1.
Oxidoreductase: NADP; Complete proteome.
BINDING: 186 186 IMP (POTENTIAL).
NP_BIND: 216 239 NADP (POTENTIAL).
SEQUENCE 349 AA; 38319 MW; D4D109B06B93072 CCK64;

Alignment Scores:
Ref. No.: 2..66 Length: 349
Core: 66.00 Matches: 26
Percent Similarity: 43.12% Conservatave: 21
Fast Local Similarity: 23.85% Mismatches: 36
Every Match: 10.05% Indels: 24
Gaps: 4

5-09-762-027-4 (1-388) x GUAQ_BUCAL (1-349)
22 TCTCAGCTCCTCTCTGCTACAAAGCTGTGATTTCTTCT-----GTAAG 69
39 TYRserAlaSerIleTrpserCylIleProlIleAlaAlaAsnMetAspThrIleCyl 58
70 ACAGAACTCTGCTGTTAACTACAGAGAACCA1AAC1TCC1C-----AAAGT 117
59 rhpheglmetValIysSerIleuSerIlePheAsnIleLeuIleAlaValnHISGly 76
118 TATGCTTAACAGTACAGACAGTACAGTACAGTATTTTATIGTTTCATCCATAGATA 177
79 TYRserPheglu-----GluTrpIlyAsnPhaValCylSerIlySerIly 95
178 TTACCTAACAGAAAGCTTCCATGCC----- 204
96 LeuAsnHisValIleValSerIleGlyThrSerAsnIleAspPheLeuIlyIleLys 115
205 -----ATACGTAGAAAGTGGCAGCAGCAGTGTCTCCTACAGGICTACAAAGACTCA 256
116 IlePheLeuLeuSerSerGluLeuIlyIleCylIleAspValAlaAsnGlyIlySer 135
256 GATCATCTCTCTGAGGAGAAAGAGCTG 285
136 GluHisIleValSerPheLeuIlySer 144

-SUL1 4
409_YEAST
P31334;
RM09_YEAST STANDARD; PRT: 269 AA.
01-JUL-1993 (Rel. 26; Created)
01-OCT-1996 (Rel. 34; Last sequence update)
15-JUL-1999 (Rel. 38; Last annotation update)
60S ribosomal protein 19, mitochondrial precursor (Yml5).
MRP15 OR YGR220C OR G8520.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
11
SEQUENCE FROM N.A., AND SEQUENCE OF 20-53; 176-182 AND 257-261.
STRAIN=07173;
MEDLINE=92283262; PubMed=1597181;
Grack H.-R., Grothmann U., Kitakawa M., Scheifer K.L., Kruft V.:
"Yml5, a nucleoside-encoded mitochondrial ribosomal protein of yeast, is
homologous to L3 ribosomal proteins from all natural kingdoms and
photosynthetic organelles".
Eur. J. Biochem. 206:373-380(1992).
[2]
SEQUENCE FROM N.A.
STRAIN=S288C;
MEDLINE=96267763; PubMed=8701610;
van der Aart Q.J.M., Kleine K., Steensma H.Y.;

```

```

*See more analysis of the 4-nb USM1-YIM5-PEP15-DIE2-SM11-Phe1-YIM5-
PEP15 from the right arm of Saccharomyces cerevisiae chromosome
VII.*
1. Yeast (20:85-99(1992)).
2. PRELIMINARY PARTIAL SEQUENCE.
3. MEDLINE 997623; JOURNAL 1997;
4. Grack H.-R., Grothmann U., Kitakawa M., Scheifer K.L., Kruft V.:
5. "Yml5, a nucleoside-encoded mitochondrial ribosomal protein of yeast,
6. and N-terminal sequencing".
7. FEBS Lett. 242:1-8(1992).
8. FUNCTIONAL COMPONENT OF THE LARGE SUBUNIT OF MITOCHONDRIAL
9. RIBOSOME.
10. SUBCELLULAR LOCATION: Mitochondrial.
11. SIMILARITY: BELONGS TO THE L3P FAMILY OF RIBOSOMAL PROTEINS.
12.
13. This Swiss-Prot entry is copyright, it is produced through a collaboration
14. between the Swiss Institute of Bioinformatics and the EMBL outstation at
15. the European Bioinformatics Institute. There are no restrictions on its
16. use by non-profit institutions as long as its content is in no way
17. modified and this statement is not removed. Usage by and for commercial
18. entities requires a license agreement (See http://www.isb-sib.ch/announce/
19. or send an email to license@isb-sib.ch).
20.
21. EMBL X52447; CAA49146.1;
22. EMBL A27411; CAA49146.1;
23. EMBL X50047; CAA49146.1;
24. PIR: S44551; REB137.
25. PIR: S44551; S26337.
26. SIO: S00452; MRP15.
27. InterPro: IPR000977; Ribosomal_L3.
28. Pfam: PF002977; Ribosomal_L3_1.
29. PROSITE: PS001374; Ribosomal_L3_1.
30. PROSITE: PS00174; Ribosomal_L3_1.
31. RN: Ribosomal protein, mitochondrial; Transic peptide.
32. TRANSIT 1 19 MITOCHONDRION.
33. CHAIN 20 279 60S RIBOSOMAL PROTEIN L5.
34. FT CONFLICT 135 139 E - V (IN REF. 1).
35. SEQUENCE 269 AA; 29790 MW; AWA41FA5CE9337E5 CCK64;

Alignment Scores:
Ref. No.: 4..11 Length: 269
Core: 64.50 Matches: 23
Percent Similarity: 47.95% Conservatave: 12
Fast Local Similarity: 31.51% Mismatches: 23
Every Match: 9.63% Indels: 15
Gaps: 4

US-9-762-027-4 (1-388) x Rhoe_Yeast (1-259)
269 GTCCTTCCCTCAAGAAAGATGA TGAATTTCTTTTAGACCTGTGAGGACAGCTGCT 224
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
136 AlaAlaPheAlaValIleAspIleGlySerIleLeuIle-----ProGlyIleIleLeu 155
233 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 194
156 L.SprosserPheIleIlySerGluGluGluGluValIleValIleValIleValIle 175
199 GITTCTTTTACATATATATATATATATATATATATATATATATATATATATATAT 137
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
176 rhpheglmetValIleLeuIlySerIleGlyIlyIlyIlyIlyIlyIlyIlyIlyIly 193
196 rhpheglmetValIleLeuIlySerIleGlyIlyIlyIlyIlyIlyIlyIlyIlyIly 96
134 GGTTrpSerIleMetHisAlaGlnHisGlyCylSerIlyCyl 206

RESULT 5
US-9-762-027-4 (1-388) x Rhoe_Yeast (1-259)
1. US20_YEAST STANDARD; PRT: 215 AA.
2. US20_YEAST
3. 15-OCT-2001 (Rel. 40; Created)
4. 15-OCT-2001 (Rel. 40; Last sequence update)
5. 15-OCT-2001 (Rel. 40; Last annotation update)

```











```

D1 01-NOV-1997 (Rel. 35, Last sequence update)
D1 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein HP0269.
GN HP0269.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group.
OC Helicobacter.
OC NCBI_TaxID=210.
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RA MEDLINE=97394167; PubMed 9252185;
RA Tomb J.F., White O., Kervatave A.R., Clayton K.A., Sutton G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Khadelak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT *The complete genome sequence of the gastric pathogen Helicobacter
RI pylori*.
RL Nature 388:539-547(1997).
CC -1- SIMILARITY: BELONGS TO THE UPF0004 FAMILY.
CC .....
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC .....
CC ENBL: AEO00546; AAD07337.1;
DR TIGR: HP0269;
DR InterPro: IPR001861; UPF0004.
DR Pfam: PF00915; UPF0004; 1.
DR TIGRFAMs: TIGR00089; UPF0004; 1.
DR PROSITE: PS01278; UPF0004; 1.
RW Hypothetical protein: Complete proteome.
SQ SEQUENCE 437 AA: 49423 MW: 38CD466BC1CECF CRC64:

```

```

Alignment Scores:
Pred. Ko.: 5.04 Length: 437
Score: 62.00 Matches: 24
Percent Similarity: 52.00% Conservative: 15
Best Local Similarity: 32.00% Mismatches: 32
Query Match: 9.44% Indels: 4
DB: 1 Gaps: 2

```

US-09-762-027-4 (1-388) x Y269\_HELPV (1-437)

```

QY 163 TCAGTCGGTAGAGTATTAGCTACAGAAACCTTC--ATTCGATACAGAGAAATCG 219
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 93 SerValSerPheValLeuGlyAlaArgAsnValSerLysIleSerGlnValIleHisGly 112
QY 220 CAGGAGCAGTGTGCTACAGAGCTCTACAAAGAACTTTCATATCTTTTCAAGGAAAG 17
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 113 GluLysAlaValGluValAlaIleAspTyrAspLysSerAlaValIleAspIlePhe 112
QY 280 AAGCTGAGAGCTACATACATAGCTGTGCTTCTTACATCTC-----AGAAATCG 33
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 133 GluLysLysAlaGlnIleAlaGlySerLeuAsnIleSerIleGlyAspLysLysGly 154
QY 331 AGATCTGTATAATCCCTTAGAAAGCAATCCCGATGATCAATA 375
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 153 AlaTyrCysIleValPronHisThrArgGlyLysGlnIleSerIle 167

```

RESULT 12  
GCH2\_HELPV  
ID GCH2\_HELPV STANDARD: PRF: 192 AA.  
AC G9ZLA42;

```

D1 15-JUL-1996 (Rel. 36, Last sequence update)
D1 15-JUL-1996 (Rel. 36, Last sequence update)
DE Influenza C virus (strain C/Yamagata/10/81).
DE Influenza C virus (strain C/Yamagata/10/81).
DE Viruses: ssRNA negative-strand viruses; Orthomyxoviridae;
GN NS.
OS Influenza C virus (strain C/Yamagata/10/81).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group.
OC Helicobacter.
OC NCBI_TaxID=210.
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RA MEDLINE=97394167; PubMed 9252185;
RA Tomb J.F., White O., Kervatave A.R., Clayton K.A., Sutton G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Khadelak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldrum J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT *The complete genome sequence of the gastric pathogen Helicobacter
RI pylori*.
RL Nature 388:539-547(1997).
CC -1- SIMILARITY: BELONGS TO THE GIP CYCLOHYDROLASE II FAMILY.
CC .....
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC .....
CC ENBL: AEO00546; AAD06324.1;
DR TIGR: HP0269;
DR InterPro: IPR000926; GIP_Cyclohydrol2.
DR Pfam: PF00925; GIP_Cyclohydrol2; 1.
DR TIGRFAMs: TIGR00050; GIP2; 1.
RW Ribonuclease: Complete proteome.
SQ SEQUENCE 192 AA: 21678 MW: 6ABHYDECC8F640 CRC64:

```

```

Alignment Scores:
Pred. Ko.: 5.95 Length: 192
Score: 61.50 Matches: 19
Percent Similarity: 46.43% Conservative: 7
Best Local Similarity: 33.93% Mismatches: 21
Query Match: 9.46% Indels: 9
DB: 1 Gaps: 2

```

US-09-762-027-4 (1-388) x GCH2\_HELPV (1-192)

```

QY 139 GAGTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 196
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 135 GluLeuLeuAlaTyrTyrArgGlyLeuLysLysMetArgLeuLeuHisMetArgLeuLys 154
QY 199 A. SerValSerPheValLeuGlyAlaArgAsnValSerLysIleSerGlnValIleHisGly 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 153 GluLysAlaValGluValAlaIleAspTyrAspLysSerAlaValIleAspIlePhe 173
QY 441 GCTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D0 174 -----AGAGTATATATATATATATATATATATATATATATATATATATATATATAT 167

```

RESULT 13  
VNSI\_HELPV  
ID VNSI\_HELPV STANDARD: PRF: 286 AA.  
AC P06895; VNSI\_HELPV; G9ZLA42;  
D1 01-JAN-1986 (Rel. 05, Last sequence update)  
D1 01-JAN-1986 (Rel. 05, Last sequence update)  
D1 15-JUL-1996 (Rel. 36, Last annotation update)  
DE Hypothetical protein: Complete proteome.  
GN NS.  
OS Influenza C virus (strain C/Yamagata/10/81).  
OC Viruses: ssRNA negative-strand viruses; Orthomyxoviridae;





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# 3M nucleic - protein search, using frame\_plus\_n2p model

Run on: January 2, 2003, 14:21:49 ; Search time 37.5 seconds  
(without alignment) 2757.399 Million cell updates/sec

File: US-09-762-027-4  
Sequence: 1 agctgcctacacagagacat.....atccataaangtalarngc 365

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delcp 6.0, Delcxt 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 161654

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=frame\_n2p\_model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09762027/runcat\_02012003\_085613\_2636/afp\_query.fasta\_1.563  
-DB-A=Geneseq\_101002 -OPMT=fastan -SUFFIX=n2p\_rag -MINMATCH=0.1 -LOCPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human:40.csl  
-LIST=45 -POCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09762027 -CGCN\_1\_L122=funat\_02012003\_085613\_2636 -NCPU=9 -ICPU=3  
-NO\_XLPEV -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -CUT=TIMEOUT=10  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database:

1: A=Geneseq\_101002:  
2: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1957.DAT  
3: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1958.DAT  
4: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1959.DAT  
5: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1960.DAT  
6: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1961.DAT  
7: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1962.DAT  
8: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1963.DAT  
9: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1964.DAT  
10: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1965.DAT  
11: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1966.DAT  
12: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1967.DAT  
13: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1968.DAT  
14: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1969.DAT  
15: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1970.DAT  
16: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1971.DAT  
17: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1972.DAT  
18: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1973.DAT  
19: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1974.DAT  
20: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1975.DAT  
21: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1976.DAT  
22: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1977.DAT  
23: /SIDS2/gcgdata/geneseq/geneseq-emb/AA1978.DAT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Length	Score	Description
1	1.7	24	2.7	Human drug metabol
2	1.6	24	2.7	Human prostatic can
3	1.5	24	2.7	Human chemokine be
4	1.4	24	2.7	Human transferrin
5	1.3	24	2.7	Human transferrin
6	1.2	24	2.7	Human drug metabol
7	1.1	24	2.7	Human drug metabol
8	1.0	24	2.7	Human drug metabol
9	0.9	24	2.7	Human drug metabol
10	0.8	24	2.7	Human drug metabol
11	0.7	24	2.7	Human drug metabol
12	0.6	24	2.7	Human drug metabol
13	0.5	24	2.7	Human drug metabol
14	0.4	24	2.7	Human drug metabol
15	0.3	24	2.7	Human drug metabol
16	0.2	24	2.7	Human drug metabol
17	0.1	24	2.7	Human drug metabol
18	0.0	24	2.7	Human drug metabol
19	0.0	24	2.7	Human drug metabol
20	0.0	24	2.7	Human drug metabol
21	0.0	24	2.7	Human drug metabol
22	0.0	24	2.7	Human drug metabol
23	0.0	24	2.7	Human drug metabol
24	0.0	24	2.7	Human drug metabol
25	0.0	24	2.7	Human drug metabol
26	0.0	24	2.7	Human drug metabol
27	0.0	24	2.7	Human drug metabol
28	0.0	24	2.7	Human drug metabol
29	0.0	24	2.7	Human drug metabol
30	0.0	24	2.7	Human drug metabol
31	0.0	24	2.7	Human drug metabol
32	0.0	24	2.7	Human drug metabol
33	0.0	24	2.7	Human drug metabol
34	0.0	24	2.7	Human drug metabol
35	0.0	24	2.7	Human drug metabol
36	0.0	24	2.7	Human drug metabol
37	0.0	24	2.7	Human drug metabol
38	0.0	24	2.7	Human drug metabol
39	0.0	24	2.7	Human drug metabol
40	0.0	24	2.7	Human drug metabol
41	0.0	24	2.7	Human drug metabol
42	0.0	24	2.7	Human drug metabol
43	0.0	24	2.7	Human drug metabol
44	0.0	24	2.7	Human drug metabol
45	0.0	24	2.7	Human drug metabol

## ALIGNMENTS

1	AA05771	Human drug metabol
2	AA05771	Human drug metabol
3	AA05771	Human drug metabol
4	AA05771	Human drug metabol
5	AA05771	Human drug metabol
6	AA05771	Human drug metabol
7	AA05771	Human drug metabol
8	AA05771	Human drug metabol
9	AA05771	Human drug metabol
10	AA05771	Human drug metabol
11	AA05771	Human drug metabol
12	AA05771	Human drug metabol
13	AA05771	Human drug metabol
14	AA05771	Human drug metabol
15	AA05771	Human drug metabol
16	AA05771	Human drug metabol
17	AA05771	Human drug metabol
18	AA05771	Human drug metabol
19	AA05771	Human drug metabol
20	AA05771	Human drug metabol
21	AA05771	Human drug metabol
22	AA05771	Human drug metabol
23	AA05771	Human drug metabol
24	AA05771	Human drug metabol
25	AA05771	Human drug metabol
26	AA05771	Human drug metabol
27	AA05771	Human drug metabol
28	AA05771	Human drug metabol
29	AA05771	Human drug metabol
30	AA05771	Human drug metabol
31	AA05771	Human drug metabol
32	AA05771	Human drug metabol
33	AA05771	Human drug metabol
34	AA05771	Human drug metabol
35	AA05771	Human drug metabol
36	AA05771	Human drug metabol
37	AA05771	Human drug metabol
38	AA05771	Human drug metabol
39	AA05771	Human drug metabol
40	AA05771	Human drug metabol
41	AA05771	Human drug metabol
42	AA05771	Human drug metabol
43	AA05771	Human drug metabol
44	AA05771	Human drug metabol
45	AA05771	Human drug metabol

```

XX      08-FEB-2001: 2001WC-US041423.
XX      11-FEB-2000: 2000US-0181856.
XX      17-FEB-2000: 2000US-0183684.
XX      25-FEB-2000: 2000US-0185141.
XX      03-MAR-2000: 2000US-0186818.
XX      09-MAR-2000: 2000US-0188345.
XX      17-MAR-2000: 2000US-0189997.
XX      * (INCYTE) INCYTE GENOMICS INC.
XX      Feng YT, Yue H, Baughn MR, Yao MG, Bandman O, Azimzai Y, Lai P;
XX      Uchishi AR, Ring HZ, Shih LL, Yang J, Pollock JL;
XX      N-PDOB: AAH6198.
XX      WPI: 2001-514673/56.
XX      Isolated polypeptide encoding a drug metabolizing enzyme useful for the
XX      diagnosis, treatment, and prevention of autoimmune/inflammatory, cell
XX      proliferative, developmental and endocrine disorders .
XX      Claim 1, Page 130: 150pp; English.
XX      The invention provides human drug metabolizing enzymes (DME) and
XX      polynucleotides encoding the DMEs. The DME can be expressed by standard
XX      recombinant methodology. DMEs and their agonists and antagonists are
XX      useful for the diagnosis, treatment, and prevention of autoimmune/
XX      inflammatory, cell proliferative, developmental, endocrine such as
XX      aneurysm, eye, metabolic, and gastrointestinal disorders, including liver
XX      disorders and infection. The present sequence represents a human DME.
XX      Sequence       294 AA:
XX
XX      Alignment Scores:
XX      Aligned NO.:          2:25e-07           Length:         294
XX      Score:              117.00             Matches:            27
XX      Percent Similarity:   53.10%             Conserves:          0
XX      Best Local Similarity: 93.10%             Mismatches:        1
XX      Query Match:         17.81%             Indels:            1
XX      Gaps:                22                 Gaps:              0
XX
XX      JS-07-762-027-4 (1-368) x AAB56487 (1-244)
XX
XX      300 ATGCTGTGTCATCACTCTCAGAAACCTCCAGATTCTGTATAAATCGTAAAGAAGAGC 359
XX      |||||||
XX      1 MetLeuValLeuHisAsnSerGln,LysGluHisIleLeuPheLysSerLeuLeuLysSer 40
XX
XX      360 ATCCCTGAATCCATAAANGATATAT 364
XX      ||| |||||
XX      21 Ile-ProGluSerIleLysValIlyr 28
XX
XX      RESULT 2
XX      AAB56487
XX      DE AAB56487 standard; protein; 241 AA.
XX
XX      Human prostate cancer antigen protein sequence SEQ ID No:1005.
XX
XX      Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX      neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
XX      vulniferous; gastrointestinal; nephrotoxic; antineutrophic; gynaecological;
XX      antibacterial; gene therapy; neural; immune; reproductive; renal;
XX      gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX      wound; infectious disease.
XX
XX      Homo sapiens.
XX
XX      WO200005174-A1.

```

[illegible]

```

11 key Location/Qualifiers
12 Peptide 1..16
13 /label= Signal_peptide
14 /label= 17..136
15 /label= chemokine_beta-1-
16
17
18 W59635520-A1.
19
20 12-DEC-1996.
21
22 uc-JUN-1995; 95MO-US07171.
23
24 uc-JUN-1995; 95MO-US07171.
25
26 (MIM#-) HUMAN GENOME SCI INC.
27 (SMR-) SMITHKLINE BEECHAM CORP.
28
29
30 11 4. McCreary JL.
31
32 Wt1: 1997-043142/04.
33 WtPSDB: AAT49221.
34
35
36 Isolated human chemokine beta-12 and corresponding cDNA sequence
37 used to develop probes for treating e.g. tumours, auto-immune
38 diseases and inflammatory and infective diseases
39
40
41 Claim 1: Page 43; 55pp; English.
42
43
44 The present sequence represents the polypeptide comprising amino acids
45 16 to 120 chemokine beta-12 (Ckbeta-12). Ckbeta-12 can be used for the
46 diagnosis and treatment of conditions associated with expression of the
47 polypeptide. In particular, it can be used to inhibit the proliferation
48 of haematopoietic cells such as bone marrow stem cells. It can be used
49 for adjunct protective treatment during cancer chemotherapy or to induce
50 proliferation of leukaemic cells. It can also be used e.g. to treat
51 psoriasis, to treat solid tumours, to enhance host defences against
52 resistant chronic and acute infections, to treat T-cell mediated
53 autoimmune diseases, to stimulate wound healing and prevent scarring;
54 to treat fibrotic disorders such as liver cirrhosis, cystic fibrosis and
55 pulmonary fibrosis, to regulate haematopoiesis, to treat sinusitis and
56 rhinitis, or to treat cancers. Ckbeta-12 analogues,
57 trichostatin or acetalatis, or to treat cancers. Ckbeta-12 analogues
58 can be used e.g. in the treatment of certain autoimmune diseases
59 like multiple sclerosis, to treat allergic reactions, psoriasis,
60 leukaemia, bone marrow failure, silicosis, rheumatoid arthritis, and
61 eosinophilic syndrome.
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930

```

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

CC occurring genomic sequences. HIFs, and its catalytic or immunoregulatory
CC fragments are useful for screening libraries of compounds in a variety of
CC drug screening techniques. Antibodies which specifically bind HIFs may be
CC used for the diagnosis of disorders associated with the expression of
CC HIFs, or in assays to monitor patients being treated with HIFs or
CC agonists, antagonists or inhibitors of HIFs. The present sequence
CC represents an HIFs protein of the invention.
XX
XX
XX
Sequence 302 AA:
XX
XX
XX
Alignment Scores:
XX
XX
XX
Precl. No.: 0 105
XX
XX
XX
Score: 77.00
XX
XX
XX
Percent Similarity: 79.31%
XX
XX
XX
Best Local Similarity: 62.07%
XX
XX
XX
Query Match: 11.72%
XX
XX
XX
DB: 22
XX
XX
XX
Gaps: 0
XX
XX
XX
US-09-762-027-4 (1-388) x AAB73512 (1-302)
XX
XX
XX
QY 300 ATGCTTGCTTCATACCTCCAGACCTCAGATCTGTATTAATCTTAAAGAAAGAG 352
XX
XX
XX
E0 1 MettelleububashanSerHislyshuLeuAlaLeuTyrlySerSerAlaAlaSer 20
XX
XX
XX
QY 360 ATCCCTGATTCATTAANGTATAT 364
XX
XX
XX
E0 21 lle-ProGluSerLeuLysValTyr 28
XX
XX
XX
RESULT 5
XX
XX
XX
AAB97238
XX
XX
XX
ID ABB97238 standard; Protein: 302 AA.
XX
XX
XX
AC ABB97238;
XX
XX
XX
DI 27-JUN-2002 (first entry)
XX
XX
XX
DE Novel human protein SEQ ID NO: 506.
XX
XX
XX
Human; antihaemic; vulnerary; antiinflammatory; immunomodulatory;
XX
XX
XX
antiinfectivity; cerebroprotective; cytosstatic; rheumatic; gene therapy;
XX
XX
XX
neuroprotective; antiparkinsonian; protein therapy; ESI;
XX
XX
XX
expressed sequence tag.
XX
XX
XX
Hom0 sapiens.
XX
XX
XX
FN WO200222660-A2.
XX
XX
XX
PD 21-MAR-2002.
XX
XX
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
XX
XX
PK 11-SEP-2000; 2000US-0659671.
XX
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Q, Fan F:
XX
XX
XX
Xue AJ, Yang Y, Wehrman T, Urmanac RT:
XX
XX
XX
WPI: 2002-292408/33.
XX
XX
XX
DR N-PSDB: ABBN32424.
XX
XX
XX
PT An isolated polynucleotide for treating diseases associated with its
XX
XX
XX
encoded polypeptide such as cancer and multiple sclerosis -
XX
XX
XX
Example 2; SEQ ID NO 506; 509pp; English.
XX
XX
XX
The present invention provides the protein and coding sequence of the
XX
XX
XX
novel human proteins. These were isolated from expressed sequences tags
XX
XX
XX
(ESTs). They can be used to stimulate cell growth, to regulate
XX
XX
XX
haematopoiesis e.g. to treat aplastic anaemia, to help tissue repair,
XX
XX
XX
e.g. in burn treatment, to regulate the immune system e.g. to treat
XX
XX
XX
multiple sclerosis, to regulate actinin or inhibit e.g. to treat
XX
XX
XX
infertility, to regulate haemostasis or thrombolysis e.g. to treat

```

[illegible]

CC The invention provides human drug metabolizing enzyme (DME) polypeptides  
 CC and polynucleotides. DMEs can be expressed by standard recombinant  
 CC methodology. The DME polypeptides, polynucleotides, modulators and  
 CC antibodies are useful from diagnosing, treating and preventing autoimmu-  
 CC /inflammatory (e.g., allergies, anemia, asthma, osteoporosis, rheumatoid  
 CC arthritis, atopic dermatitis, glomerulonephritis and irritable bowel  
 CC syndrome, trauma, and bacterial, viral, parasitic, protozoal, fungal,  
 CC helminthic infections), cell proliferative (e.g., cancer, atherosclerosis,  
 CC and hepatitis), developmental (e.g., cerebral palsy and retardation),  
 CC endocrine (e.g., thrombosis, goiter, hyperplasia, amenorrhea and  
 CC dysmaturia), eye (e.g., iritis and glaucoma, metabolic (e.g., Addison's  
 CC disease, diabetes and obesity), and gastrointestinal disorders (e.g.,  
 CC constipation, nausea, abdominal angina, ulcerative colitis, diarrhea and  
 CC Crohn's disease). The DME polypeptides are useful in drug screening  
 CC techniques, and to analyse the proteome of a tissue or cell type. The DME  
 CC polynucleotides are useful for creating knock-in humanized animals or  
 CC transgenic animals to model human diseases, and in somatic or germ-line  
 CC gene therapy. The present sequence represents a human DME polypeptide  
 CC sequence.

XX Sequence 302 AA:

Alignment Scores:  
 Pred. No.: 0.559 Length: 302  
 Score: 72.00 Matches: 16  
 Percent Similarity: 82.764 Conservative: 6  
 Best Local Similarity: 55.178 Mismatches: 4  
 Query Match: 10.968 Indels: 1  
 DB: 23 Gaps: 0

US-09-762-027-4 (1-388) x ABB07518 (1-302)

QY 300 ATGCTGCTTCATACCTCAGACAGCTGCAGATTCGTGTAATAATCCTTAAGAAAGAGC  
 Db 1 MetleuleuInsnaSersGluArgLeuAlaLeuPhelySerLeuAlaArgSer 29  
 21 Ile-ProGluSerleuInsValTyr 28

RESULT 7  
 AAE48010 ID AAB48010 standard; Protein; 327 AA.  
 XX AAB48010:  
 AC AAB48010:  
 DI 15-MAR-2001 (first entry)  
 XX  
 DE Rat OX2R polypeptide sequence.  
 XX  
 KW OX2R Protein; OX2RH1; OX2RH2; OX2RH3; OX2RH4; OX2RH1.2; Ischaemia;  
 KW antiinflammatory; cytosolic; neuroprotective; neurotrophic; rodent;  
 KW antiarteriosclerotic; vasodilator; immunosuppressive; antirheumatic;  
 KW antiarthritic; gene therapy; rat.  
 XX  
 OS Rattus rattus.  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 1..24  
 FI /note= "signal peptide"  
 FI Protein 25..327  
 FT /note= "mature protein"  
 XX  
 MO200070045-A1.  
 XX  
 PN 23-NOV-2000.  
 XX  
 PD 11-MAY-2000; 2000WO-US12998.  
 XX  
 PR 13-MAY-1999; 99GB-0011123.  
 PR 03-NOV-1999; 99GB-0025989.  
 XX  
 PA (MED1-) MEDICAL RES COUNCIL.

PA (SHE) Schematic Corp.  
 XX  
 XX  
 P1 Barclay AN, Brown MR, Gorman DM, Lahier LL, Wright JJ;  
 P1 Chelwinski R, Phillips RH, Beck RM, Sedgwick JR;  
 XX  
 XX  
 P1 WPI: 2001-01-27/02;  
 P1 N-PSID: AAB48010, AAB48010.  
 XX  
 P1 Mammalian OX2R proteins and DNA sequences useful for modulating the  
 P1 physiology and development of a cell -  
 P1  
 P1 Claim 1: Page 21-22; 14pp; English.

XX The invention relates to rodent or primate OX2R proteins, especially  
 CC OX2RH1, OX2RH2, OX2RH3, OX2RH4, or OX2RH1.2 proteins. Agonists and  
 CC antagonists of the OX2RH sequences can be used to modulate physiology  
 CC or development of a cell, particularly for enhancing myeloid function or  
 CC enhancing immunity. The sequences can be used to identify non-OX2 ligands  
 CC for an OX2R. The polypeptides and polynucleotides can be used to treat  
 CC inflammatory, leukoproliferative, neurodegenerative or post-traumatic  
 CC conditions, including atherosclerosis, multiple sclerosis, ischaemia,  
 CC neurodegeneration, rheumatoid arthritis, and autoimmunity. The present  
 CC sequence represents the rat OX2R polypeptide.

XX Sequence 27 AA:

Alignment Scores:  
 Pred. No.: 1.9 Length: 27  
 Score: 66.50 Matches: 17  
 Percent Similarity: 44.44 Conservative: 15  
 Best Local Similarity: 23.61 Mismatches: 35  
 Query Match: 10.434 Indels: 5  
 DB: 22 Gaps: 1

US-09-762-027-4 (1-388) x AAB48010 (1-327)

QY 154 TATGAGTTCAGTCCGATAGATATAGCTAACGAGAACCTTCCATGCCCACTAGACA 213  
 Db 212 PhocysValValSerHisLeuThrThiGlyAsnGlnSerLeuSerIleGluGlyArg 231  
 QY 214 AACCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 273  
 Db 232 GlyGlySerIleLeuLeuGlySerIleIle-----GlnIleIleIlePro 246  
 QY 274 GAAAG 333  
 Db 247 SerIleIleLeuIleIleIleGlyGlyIleGlyLeuLeuIleSerGlyGlyArg 266  
 QY 334 TCTGTATAGATCCTTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 365  
 Db 267 LysGlySerLeuProGlySerValAlaIleProCasp 276

RESULT 8  
 ABB63656 ID ABB63656 standard; Protein; 1229 AA.  
 XX ABB63656:  
 AC ABB63656:  
 DI 26-MAR-2002 (first entry)  
 XX  
 DE Biosynthetic development of polypeptide SEQ ID NO 1776.  
 XX  
 FH Biosynthetic development of polypeptide; cell signalling; insecticide;  
 FH Pharmacological  
 KW Dirosophila melanogaster  
 XX  
 OS Dirosophila melanogaster  
 XX  
 PN W020017042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PR 23-MAR-2001; 2001WO-US09231.  
 XX



```

XX      23-MAR-2000; 20000US-191637P.
FR      11-JUL-2000; 20000US-061415U.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW:
XX      WPI: 2001-656860/75.
DR      N-PSDB: ABL07759.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
XX      Disclosure: SEQ ID NO 17760; 21pp + Sequence Listing: English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention
XX      discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
XX      (ABB57737-ABB72072).
CC      The sequence data for this patent did not form part of the prior art
CC      specification, but was obtained in electronic format directly from WIP
CC      at ftp.wipo.int/pub/published_pat_sequences.
XX
XX      Sequence 1229 AA:
SU
Alignment Scores:
Pred. No.:          2.84          Length:          1229
Score:              68.50          Matches:         22
Percent Similarity: 43.86%        Conservative:    26
Best Local Similarity: 19.30%      Mismatches:     39
Query Match:        10.43%         Indels:         25
DB:                 22             Gaps:           4
US-05-762-027-4 (1-388) x ABB63656 (1-1229)
QY      13 CGGACATATTCTCAGTCGTCCTCTCGTGTACAAAGCCCTGGATTTCTTGCTCATGGAGA 72
Db      ||||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      122 Argasnillepe-----LeulePhualeueroitrrilleleuthrPeVal----- 137
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      73 GAACCTCTGGTTTAATCTACAGAACCCCAACTTCCTGAGAGCTTAGTGAATACAGTGA 132
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      138 -----Alserelyleuilelesnilyrllearkyrfhelrnel----- 143
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      133 CAACGTGAGTCAGTTGATTTATIGITTCAGTCGTAGAAATAAAT----- 163
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      150 -----VallieghlnishPhecysilesesasnrrpargalalelleuleuerydy 164
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      184 -----AACAGAACCTTCGATTCCTCCCTACAGCAAAACCTGGAG 242
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      167 gluteluclnglyleuPhetrprhserrlserlahtlrleullePerlofrrtprser 187
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      223 CAGCGAGTCGTGCCCTCAGAGCTTACAACAACACTTCAGATCAATCTTGAAGAAAAGAG 262
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      187 glinllellcglvglyValserleucluhylaslaserfnslseuserclyAsnserrary 205
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      283 CTGCAATGCTGCATACAGANTGCTGTCTCTCACTAAGCTGGA 324
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      207 ValTYrSerAlaValAsnIlellelleuTYrThrretary 220
Db      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX      RESULT 9
XX      ABB62231
XX      ID ABB62231 standard; Protein: 2441 AA.
XX      AC ABB62231:
XX
XX      26-MAR-2002 (first entry)
XX
XX      Drosophila melanogaster polypeptide SEU ID NO 1-1405.

```

```

XX XX Insecticide; development; cell signalling; insecticide;
XX XX Protein;
XX XX Insecticide;
XX XX W04001704-2-Aa;
XX XX 27-SEP-2001;
XX XX 23-MAR-2001; 23-MAR-2001;
XX XX PR 23-MAR-2001; 23-MAR-2001;
XX XX PK 11-JUL-2001; 11-JUL-2001;
XX XX PA (PDB ) PDB ID: 1A;
XX XX Vector JD; Adams M; Chishti, Myers EW;
XX XX WP1: 2001-05267-975;
XX XX DR N-Psiab; ARI06714;
XX XX PS New isolated protein; detection reagent for detecting ludo or more
XX XX PT genes from insects not for elucidating cell signalling and cell-cell
XX XX interactions ;
XX XX Disclosure Sheet No. 2145; 2145 - Sequence listing; English.
XX XX The invention relates to an isolated nucleic acid detection reagent
XX XX capable of identifying one or more genes from Drosophila. The invention
XX XX useful in determining the biology and in elucidating cell signalling and
XX XX cell-cell interaction in higher eukaryotes for the development of
XX XX insecticides, therapeutic and pharmaceutical drugs. The invention
XX XX discloses genomic DNA sequences (AB116176-AB116511), expressed cDNA
XX XX sequences (AB116170-AB116175) and the encoded proteins
XX XX CC (AA857737-AB857872).
XX XX CC The sequence data for this patent did not form part of the printed
XX XX CC specification, but was obtained in electronic format directly from WIPO
XX XX at ftp://ipo.int/pub/PUBLISHED_PCT_sequences.
XX XX Sequences 214; AA:
XX XX
XX XX Alignment Scores:
XX XX Pred. NO.: Length: 2141
XX XX Score: 45 Matches: 15
XX XX Percent Similarity: 4.4 Conservative: 6
XX XX Best Local Similarity: 4.4 Mismatches: 11
XX XX Query Match: 1.745 Indels: 3
XX XX DB: Gaps: 1
XX XX
XX XX US-07-752-074 (1-1755) X (Amo221) (1-2141)
XX XX
XX XX 23 AGAGACGATCAATTCATTTTGAAGAAGCAGTGGG 293
XX XX 111111111111111111111111111111111111
XX XX 2200 SerProSerLeuLeuSerLeuAsnHisAspIleArglyShePrCaGPrLLIGclY 2279
XX XX
XX XX 232 ACCTGTGCTTTTCCATTAATTCAGATGCGAATGAAAGGCTTC 103
XX XX 111111111111111111111111111111111111
XX XX ED 2200 ThrGlnArgTrpAlasL L-----LProValTrpAsnSnpW 2291
XX XX
XX XX RESULT 10
XX XX AB854470
XX XX ID AB854470 standard protein for AA.
XX XX
XX XX AA AB854470:
XX XX
XX XX 23-MAR-2001; 23-MAR-2001;
XX XX LE ProSophia medInoclastin B peptide Spg ID NO 20202.
XX XX Inositol; developmental biology; cell signalling; insecticide;
XX XX pharmaceutical.
XX XX

```

[illegible]

1 2324141  
 2 2324141  
 3 2324141  
 4 2324141  
 5 2324141  
 6 2324141  
 7 2324141  
 8 2324141  
 9 2324141  
 10 2324141  
 11 2324141  
 12 2324141  
 13 2324141  
 14 2324141  
 15 2324141  
 16 2324141  
 17 2324141  
 18 2324141  
 19 2324141  
 20 2324141  
 21 2324141  
 22 2324141  
 23 2324141  
 24 2324141  
 25 2324141  
 26 2324141  
 27 2324141  
 28 2324141  
 29 2324141  
 30 2324141  
 31 2324141  
 32 2324141  
 33 2324141  
 34 2324141  
 35 2324141  
 36 2324141  
 37 2324141  
 38 2324141  
 39 2324141  
 40 2324141  
 41 2324141  
 42 2324141  
 43 2324141  
 44 2324141  
 45 2324141  
 46 2324141  
 47 2324141  
 48 2324141  
 49 2324141  
 50 2324141  
 51 2324141  
 52 2324141  
 53 2324141  
 54 2324141  
 55 2324141  
 56 2324141  
 57 2324141  
 58 2324141  
 59 2324141  
 60 2324141  
 61 2324141  
 62 2324141  
 63 2324141  
 64 2324141  
 65 2324141  
 66 2324141  
 67 2324141  
 68 2324141  
 69 2324141  
 70 2324141  
 71 2324141  
 72 2324141  
 73 2324141  
 74 2324141  
 75 2324141  
 76 2324141  
 77 2324141  
 78 2324141  
 79 2324141  
 80 2324141  
 81 2324141  
 82 2324141  
 83 2324141  
 84 2324141  
 85 2324141  
 86 2324141  
 87 2324141  
 88 2324141  
 89 2324141  
 90 2324141  
 91 2324141  
 92 2324141  
 93 2324141  
 94 2324141  
 95 2324141  
 96 2324141  
 97 2324141  
 98 2324141  
 99 2324141  
 100 2324141

FD	30	GlnAlaIysValSerSerMetCHeLeProLeuGlnIcInlyAlaIsIndeLr.LinDeLr.Lr.
Seq	359	AATAATCGTTAGAAAGACCAICCCG 365
		:
ED	50	GlySerLeuArgLySerLeuPro 58
RESULT 1:		
ID	Ab013939	standard: Protein; 217 AA.
XX	Ab013939;	
XX	19-FEB-2002	(first entry)
DI	XX	
EE	XX	Human human diagnostic protein #13930.
KM	XX	Name: Chromosome mapping; gene mapping; gene therapy; forensics;
XX	XX	tissue supplement; medical imaging; diagnosis; genetic disorder;
XX	XX	Homo sapiens.
XX	XX	1-20175067-A2.
PC	XX	1-NOV-2001.
FE	XX	20-MAR-2001; 2001MO-US08631.
FR	XX	31-MAR-2000; 2000US-0540217.
XX	XX	2--FUG-2000; 2000US-0649167.
PA	XX	(HWSE-) HWSEQ INC.
XX	XX	
E1	XX	Lermanac RT, Liu C, Tang YT;
LR	XX	Wnt: 2001-639362/73.
R	XX	FSCB; AAS78126.
FI	XX	New isolated polynucleotide and encoded polypeptides, useful in
PI	XX	diagnostics, forensics, gene mapping, identification of mutations
FT	XX	responsible for genetic disorders or other traits and to assess
FT	XX	biocompatibility -
ES	XX	Claim 20: SEQ ID NO 44296; 103pp; English.
XX	XX	
CC	XX	The invention relates to isolated polynucleotide (I) and
CC	XX	polypeptide (II) sequences. (I) is useful as hybridisation probes
CC	XX	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	XX	and gene mapping, and in recombinant production of (II). The
CC	XX	polynucleotides are also used in diagnostics as expression tags
CC	XX	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	XX	to restore normal activity of (II) or to treat disease states involving
CC	XX	(II). (II) is useful for generating antibodies against it, detecting
CC	XX	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	XX	a food supplement. (II) and its binding partners are used in treating,
CC	XX	healing or sites expressing (II). (I) and (II) are useful for treating
CC	XX	disorders involving aberrant protein expression or abnormal activity.
CC	XX	The polypeptide and polynucleotide sequences have applications in
CC	XX	diagnostics, forensics, gene mapping, identification of mutations
CC	XX	responsible for genetic disorders or other traits to assess biocompatibility
CC	XX	and to produce other types of data and products dependent on DNA and
CC	XX	amino acid sequences. AB000010-AB030377 represent novel R and
CC	XX	diagnostic amino acid sequences or the invention.
CC	XX	Note: The sequence data for this patent did not appear in the EMBL
CC	XX	specification, but was obtained in electronic format directly from Wf
CC	XX	at ftp.wipo.int/pub/published_pat_sequences.
SD	Sequence	217 AA:
Alignment Scores:		
Score:	6.93	Length: 217
Fred. No.:	64.00	Matches: 19
Percent Similarity:	54.00%	Conservatism: 12
Best Local Similarity:	30.00%	Mismatches: 12

Accession	Gene	Species	Length	Source
U00001.2	oriC	<i>Escherichia coli</i>	2454	GenBank
U00002.2	oriT	<i>Escherichia coli</i>	2454	GenBank
U00003.2	oriV	<i>Escherichia coli</i>	2454	GenBank
U00004.2	oriW	<i>Escherichia coli</i>	2454	GenBank
U00005.2	oriX	<i>Escherichia coli</i>	2454	GenBank
U00006.2	oriY	<i>Escherichia coli</i>	2454	GenBank
U00007.2	oriZ	<i>Escherichia coli</i>	2454	GenBank
U00008.2	oriA	<i>Escherichia coli</i>	2454	GenBank
U00009.2	oriB	<i>Escherichia coli</i>	2454	GenBank
U00010.2	oriC	<i>Escherichia coli</i>	2454	GenBank
U00011.2	oriD	<i>Escherichia coli</i>	2454	GenBank
U00012.2	oriE	<i>Escherichia coli</i>	2454	GenBank
U00013.2	oriF	<i>Escherichia coli</i>	2454	GenBank
U00014.2	oriG	<i>Escherichia coli</i>	2454	GenBank
U00015.2	oriH	<i>Escherichia coli</i>	2454	GenBank
U00016.2	oriI	<i>Escherichia coli</i>	2454	GenBank
U00017.2	oriJ	<i>Escherichia coli</i>	2454	GenBank
U00018.2	oriK	<i>Escherichia coli</i>	2454	GenBank
U00019.2	oriL	<i>Escherichia coli</i>	2454	GenBank
U00020.2	oriM	<i>Escherichia coli</i>	2454	GenBank
U00021.2	oriN	<i>Escherichia coli</i>	2454	GenBank
U00022.2	oriO	<i>Escherichia coli</i>	2454	GenBank
U00023.2	oriP	<i>Escherichia coli</i>	2454	GenBank
U00024.2	oriQ	<i>Escherichia coli</i>	2454	GenBank
U00025.2	oriR	<i>Escherichia coli</i>	2454	GenBank
U00026.2	oriS	<i>Escherichia coli</i>	2454	GenBank
U00027.2	oriT	<i>Escherichia coli</i>	2454	GenBank
U00028.2	oriU	<i>Escherichia coli</i>	2454	GenBank
U00029.2	oriV	<i>Escherichia coli</i>	2454	GenBank
U00030.2	oriW	<i>Escherichia coli</i>	2454	GenBank
U00031.2	oriX	<i>Escherichia coli</i>	2454	GenBank
U00032.2	oriY	<i>Escherichia coli</i>	2454	GenBank
U00033.2	oriZ	<i>Escherichia coli</i>	2454	GenBank
U00034.2	oriA	<i>Escherichia coli</i>	2454	GenBank
U00035.2	oriB	<i>Escherichia coli</i>	2454	GenBank
U00036.2	oriC	<i>Escherichia coli</i>	2454	GenBank
U00037.2	oriD	<i>Escherichia coli</i>	2454	GenBank
U00038.2	oriE	<i>Escherichia coli</i>	2454	GenBank
U00039.2	oriF	<i>Escherichia coli</i>	2454	GenBank
U00040.2	oriG	<i>Escherichia coli</i>	2454	GenBank
U00041.2	oriH	<i>Escherichia coli</i>	2454	GenBank
U00042.2	oriI	<i>Escherichia coli</i>	2454	GenBank
U00043.2	oriJ	<i>Escherichia coli</i>	2454	GenBank
U00044.2	oriK	<i>Escherichia coli</i>	2454	GenBank
U00045.2	oriL	<i>Escherichia coli</i>	2454	GenBank
U00046.2	oriM	<i>Escherichia coli</i>	2454	GenBank
U00047.2	oriN	<i>Escherichia coli</i>	2454	GenBank
U00048.2	oriO	<i>Escherichia coli</i>	2454	GenBank
U00049.2	oriP	<i>Escherichia coli</i>	2454	GenBank
U00050.2	oriQ	<i>Escherichia coli</i>	2454	GenBank
U00051.2	oriR	<i>Escherichia coli</i>	2454	GenBank
U00052.2	oriS	<i>Escherichia coli</i>	2454	GenBank
U00053.2	oriT	<i>Escherichia coli</i>	2454	GenBank
U00054.2	oriU	<i>Escherichia coli</i>	2454	GenBank
U00055.2	oriV	<i>Escherichia coli</i>	2454	GenBank
U00056.2	oriW	<i>Escherichia coli</i>	2454	GenBank
U00057.2	oriX	<i>Escherichia coli</i>	2454	GenBank
U00058.2	oriY	<i>Escherichia coli</i>	2454	GenBank
U00059.2	oriZ	<i>Escherichia coli</i>	2454	GenBank
U00060.2	oriA	<i>Escherichia coli</i>	2454	GenBank
U00061.2	oriB	<i>Escherichia coli</i>	2454	GenBank
U00062.2	oriC	<i>Escherichia coli</i>	2454	GenBank
U00063.2	oriD	<i>Escherichia coli</i>	2454	GenBank
U00064.2	oriE	<i>Escherichia coli</i>	2454	GenBank
U00065.2	oriF	<i>Escherichia coli</i>	2454	GenBank
U0				

see also AAR23140-79 and AAR23722

Sequence 609 AA;

Alignment Scores:	
red. No.:	9.73
Score:	64.00
Matches:	18
Conservative:	38.218
Mismatch:	23.558
Local Similarity:	9.748
Identity:	13
Gaps:	5

5-05-702-027-4 (1-388) x AAR23174 (1-609)

[illegible]

RESULT 14  
AKR23173  
AKR23173 standard; Protein; 650 AA  
X  
X AKR23173;  
X

22-OCF-1992 (first entry)  
Mutant thermostable DNA polymerase enzyme MET-1HR 204 TAF

5'-3'; exonuclease; PC  
Thermosiphó africanus.

	Location/Qualifiers
Key:	1..2
Misc-difference	/note="residues 2-203 deleted from the native sequence"

W09206200-A.

16-APR-1992

١٠٠

28-SEP-1990; 90US-0590413.  
28-SEP-1990; 90US-0590466.  
28-SEP-1990; 90US-0590490.

(CETU ) CETUS CORP

3.  
 2.  
 1.  
 0.  
 -1.  
 -2.  
 -3.  
 -4.  
 -5.  
 -6.  
 -7.  
 -8.  
 -9.  
 -10.  
 -11.  
 -12.  
 -13.  
 -14.  
 -15.  
 -16.  
 -17.  
 -18.  
 -19.  
 -20.  
 -21.  
 -22.  
 -23.  
 -24.  
 -25.  
 -26.  
 -27.  
 -28.  
 -29.  
 -30.  
 -31.  
 -32.  
 -33.  
 -34.  
 -35.  
 -36.  
 -37.  
 -38.  
 -39.  
 -40.  
 -41.  
 -42.  
 -43.  
 -44.  
 -45.  
 -46.  
 -47.  
 -48.  
 -49.  
 -50.  
 -51.  
 -52.  
 -53.  
 -54.  
 -55.  
 -56.  
 -57.  
 -58.  
 -59.  
 -60.  
 -61.  
 -62.  
 -63.  
 -64.  
 -65.  
 -66.  
 -67.  
 -68.  
 -69.  
 -70.  
 -71.  
 -72.  
 -73.  
 -74.  
 -75.  
 -76.  
 -77.  
 -78.  
 -79.  
 -80.  
 -81.  
 -82.  
 -83.  
 -84.  
 -85.  
 -86.  
 -87.  
 -88.  
 -89.  
 -90.  
 -91.  
 -92.  
 -93.  
 -94.  
 -95.  
 -96.  
 -97.  
 -98.  
 -99.  
 -100.

20  
19  
18  
17  
16  
15  
14  
13  
12  
11  
10  
9  
8  
7  
6  
5  
4  
3  
2  
1

Interstate and intrastate with altered 5'-3' exo nuclease activity. Javitt showed a reduction in the number of active sites in e.g. for segment of deletion assays.

Clark II: page 59: trip: nullish.

activity than the active enzyme. The

are useful in many recombinant DNA techniques; resp. nucleic acid amplification by *in situ* self-sustained sequence replication (SSR) and high temp. DNA sequencing. The absence of 5'-3' nuclease activity may facilitate higher sensitivity allelic discrimination in a combined polymerase ligase chain reaction (PRLC) assay. An enhanced amt. of 5-3' exonuclease activity may be desirable in enzymes used in heteroduplex assays for the amplification and detection of a target nucleic acid sequence. Mutation of the DNA encoding particular regions of the organism can be used to prepare a range of recombinant proteins having 5-3' exonuclease activity to a complete lack of activity.

See also A060231979 and A0607322.

Sequence:	790 AA:
Element Score:	
ID No:	1472
Ei:	41.00
Ident Similarity:	37.21%
Local Similarity:	37.58%
Y Match:	3.74%
I's:	5
Caps:	
Length:	690
Matches:	25
MisMatches:	16
Mismatches:	33
Indels:	43

80 100111A1C1A-200A00011A0101010  
341 .....

348 Ingrid Valaieiri

1111 : : 111  
361 Alaphonin-Lysylstir

374 ..... 255.1  
320 ..... 126.5

38 / Catalisat

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

AA23172 standard; Protein: 75

4:AR23172;







```

      9 Phcylagargargileulellemethtyrstyphallevalleuleuerset:cysllsetha 76
    CY 295 TGACGACTTCAGCTTCTTTCCCTCAAGAAGAT----- 27
          |||||||
    CY 29 Cysserthrspneaspvalvaingluuaspsarvallysllevallysahlsyryr 48
          |||||||
    CY 262 GATCGTAAGTTCTTGTTAGACCCTGAAGCACACTGCCTCCTGCCAAATCTTACATAA 103
          |||:::|||||
    CY 49 Aspallylsspietmrcys-----thPhaser---tlp 5c
          |||
    CY 202 CAATGGAAAGCTTTCTGTAACTAAIACITCAACGACTGAACACAIAAAATTCAT 118
          |||
    CY 59 Glu-----lledtcl 62
          |||
    CY 142 ACTCACGTTGTCAGCTGTTAACGATAAAGCTTCACAA 107
          |||
          ::|||
    CY 63 thrllys--AlatrpilleyshsthrThrasply 73
          |||
RESULT 5
CY4261 PRELIMINARY: PRT: 220 AA.
Ac O94261:
E1 01-MAY-1959 (TREMBLrel_10, Created)
D1 01-MAY-1959 (TREMBLrel_10, Last sequence update)
B1 01-DEC-2001 (TREMBLrel_19, Last annotation update)
LE Hypothetical 25.0 kDa protein.
GN SPBP887.12C.
OS Schizosaccharomyces pombe (Fission yeast).
US Schizosaccharomycetes
CU Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
SC Schizosaccharomycetales; Schizosaccharomycetaceae;
CN Schizosaccharomyces.
NCBI_TaxID=4896;
KH (1)
AF SEQUENCE FROM N.A.
RA SIRA1N-972H-;
RA Beck A., Reinhardt R, Lyne M., Rajandream M.A., Bartell B.C.;
AL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
FE EMBL: AL032684; CAA21757.1; -
FE Hypothetical protein.
FE SEQUENCE 220 AA: 2497 MW: c774fc4e5053ff3 CR::4;
```

```

Alignment Scores:
      1 95      length: 220
      70 00      matches: 30
      45.548     percent Similarity: 10
      35.648     best Local Similarity: 27
      10.658     query Match: 28
      3          caps: 6

05-05-762-027-4 (1-386) x 054261 (1-220)

16 AACATATCTCAGTCCTCCTCTGAGTGTACAAAGCTGTGATTTCTTATATGAAATATTA
||||| ||||||||| ||| : : : |||
135 AsnIleTyrSerValIleuLeuProSerValGlnSerIleSerSerSer-----
76 CGCTGCTTTATTCATACAGAACCCAACTCCCTGAGACCTTTATGCTTAAATGAAAT
||||| ||||||||| ||| : : : |||
151 -----LeuIleuLeuTyrAla-----
136 CGTGAATCAGCTGAATTTATATGCTTTCAGTCCCTACAGATATATGCTATATAGAAAT
||||||| ||| : : : |||||||||
157 LysGlnSer-----AspTyrArgSerSerIleGlnGlnValLeuAlaAspLeuIleSer
196 TCCATTGCC-----ATACTGGAAGATTCGACGACGGAGCGTGCTGCTAAGATTCAT
||||| ||| : : : |||
175 AlaIArgAlaIArgTrpClnuIleValIArgAsnValIArgGlnIleLeuLeuLeu-----
247 AAGAAACTTCAGATCATCTTCTTACGCAAGAAAGACCTGAACTGCATCAATAGATTTT
||||| ||||| ||||||| |||
192 ---GlnSerClyIleSerPheLeuGlnAsnGlySLeu---AlaTyrIleIleuAspLeu
307 TGC 309

```

```

111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

[illegible]









RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Arabidopsis sequencing project:  
 RL Submitted (MAR-1999) to the EMBL/Genbank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Arabidopsis sequencing project:  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DDBJ databases.  
 DR EMBL: AL045171; CAB38964.1; -  
 DR EMBL: AL161565; CAB79490.1; -  
 DR InterPro: IPR001810; F-box.  
 DR Pfam: PF00646; F-box; 1.  
 DR SMART: SM00256; FBOX; 1.  
 DR PROSITE: PSS0161; FBOX; 1.  
 DR Hypothetical protein.  
 KW SEQUENCE 431 AA; 49655 MW; 7BCD80DDA9121C40 CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 8.38 Length: 431  
 Score: 65.50 Matches: 26  
 Percent Similarity: 46.32% Conservative: 16  
 Best Local Similarity: 27.37% Mismatches: 32  
 Query Match: 9.97% Indels: 19  
 Gaps: 6  
 DB: 10  
 US-09-762-027-4 (1-388) x Q9ST00 (1-431)  
 QY 67 TGGACAGAA---CGCTGCTTTAATCTACAGAAACCACTACCTGCTGAG-----CTT 117  
 DB 257 TTPAIGLIALAARGLIALAVALTHRLYLYSERHSLSLYLLEUYSGLUHLFHR 276  
 QY 118 TATGCTTACAGTACAGACGTAAGTCAATTGATTTATGCTTTCAGTCCGAGAG 177  
 DB 277 TYRATA-----LysGlyLeuYfValCysLeuSerValSerGlnVal 290  
 QY 178 TTAAGTACAGAAACCTTCATTCATTCAGAGAACTGCACACCAAGCAATGCTTA 237  
 DB 251 MetAspProTyrAspMetIlePheHisMetLeuValAspLeuIleLysThrCysThr 310  
 QY 238 CAGGCTACAGAAAGACTACATCATCTTTCAGAGGAAAS---AACTGTAAGTCTG 294  
 DB 311 GINGLYTRTPRASPLeuLeuThrHisMetLeuGlnGlySerProLysLeuArg 328  
 QY 295 ATAGATGCTTGTCTCTATACCTCAGAAC-----TCCAGA 333  
 DB 329 -----PheLeuThrLeuThrAsnAspHisCysArg 338  
 RESULT 14  
 Q9LN44 PRELIMINARY: PRT: 639 AA.  
 AC Q9LN44:  
 DI 01-OCT-2000 (TREMblrel. 15, Created)  
 DI 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DI 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE F18014.26.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCBITaxID=3702;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Nhat S.,  
 RA Kim C., Altairi H., Bai O., Chin C., Chou J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo B., Lee J.,  
 RA Lenz C., Li J., Liu K., Liu S., Mukharasy N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;

\*Sequence submitted to Arabidopsis thaliana BAC F18014 from chromosome  
 1.  
 RL Submitted (MAY-2000) to the EMBL/Genbank/DDBJ databases.  
 OC -1- SUBCELLULAR LOCATION: NUCLEUS (BY SIMILARITY).  
 OC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.  
 DR EMBL: A025663; AAT79444.1; -  
 DR InterPro: IPR001805; MYB-DNA-binding.  
 DR InterPro: IPR001837; IF-ELF1.  
 DR Pfam: PF00170; ELF1; 1.  
 DR Pfam: PF00243; MYB-DNA-binding; 1.  
 DR SMART: SM00256; MYB2; 1.  
 DR SMART: SM00256; MYB2; 1.  
 DR PROSITE: PSS0027; MYB2; 1.  
 DR PROSITE: PSS0027; MYB2; 1.  
 KW DNA-binding; Nuclear protein.  
 KW SEQUENCE 639 AA; 71309 MW; 672E5EB4D2EDACA CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 8.59 Length: 639  
 Score: 65.50 Matches: 20  
 Percent Similarity: 55.22% Conservative: 17  
 Best Local Similarity: 29.85% Mismatches: 27  
 Query Match: 9.97% Indels: 3  
 Gaps: 1  
 DB: 10  
 US-09-762-027-4 (1-388) x Q9LN44 (1-639)  
 QY 166 GTCATGAAATATTAATACAGAAAGAAATTTTCATTCCTACAGAAAGTGGACATAG 225  
 DB 330 TTPAIGLIALAARGLIALAVALTHRLYLYSERHSLSLYLLEUYSGLUHLFHR 344  
 QY 226 GCAATGTCATACAGAGTCTACAGAAAGAAATTTTCATTCCTACAGAAAGTGGACATAG 265  
 DB 345 AlahetCys3,AspLeuSerLysHisAlaAspLeuIleTyrGlnAspGlnLeu 366  
 QY 266 AAGTGTCTATAGATGCTTGTCTCTATACCTCAGAACCTCCAGTTCTGATGAA 345  
 DB 369 Arg-ArgGlnLysAspIlePalaLeuLysGlnPheGlnSerLeuGlnIleAsnLysH 368  
 QY 345 GTTAAAGAAAGAAATG 362  
 DB 368 IndelLysHAlaVal 393  
 RESULT 15  
 Q9LN44 PRELIMINARY: PRT: 310 AA.  
 AC Q9LN44:  
 DI 01-AUG-1998 (TREMblrel. 07, Created)  
 DI 01-AUG-1998 (TREMblrel. 07, Last sequence update)  
 DI 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE starch synthase FULL (Fragment).  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Panicoideae; Anacogonaceae; Zea.  
 NC NCBITaxID=4577;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE: 981624; PubMed-950111;  
 RA Gao M., Wanet J., Steward P.S., James M.G., Myers A.M.;  
 RI "Characterization of Dull1, a maize gene coding for a novel starch  
 synthase."  
 RL Plant Cell 10:77-112(1998).  
 DR EMBL: AF023123; AAC1015.1; -  
 DR InterPro: IPR001296; Glycosyl-transferase.  
 DR Pfam: PF00534; Glycosyl-transferase; 1.  
 DR NOFITER 1  
 KW SEQUENCE 310 AA; 34465 MW; 6410223B5530515E CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 9.62 Length: 310  
 Score: 65.00 Matches: 24  
 Percent Similarity: 43.68% Conservative: 14



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Computer Ltd.

CM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 2, 2003, 15:09:29 : Search time 13.5 seconds

(without alignments)  
1691.273 Million cell updates/sec

Title: US-09-762-027-4

Perfect score: 657

Sequence: 1 agctgcacatacgaacac.....atccatcaatgataatgag jss

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Rgapop 6.0 , Rgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 26422922 residues

Total number of hits satisfying chosen parameters: 525146

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_n2p model -DEV=xlh  
-O=/cgn2\_1/USP10.spool/US09762027/runcat\_02012003\_085614\_2671/app\_query\_fast\_11553  
-DB=Issued\_Patents\_AA -QEMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -L=0.001  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum2 -TRANS=human4v.cti  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=20000000  
-USER=US09762027@cgn2\_1.1.4.jrunat.02012003\_085614\_2671 -NCPU=6 -ICPU=3  
-NO\_XMAP=NO\_XMAP -LARGEROUDRY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=15  
-MARK\_TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6 -YGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_AA:  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfill1a1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being plotted;  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	12.6	136	5	PCT-US95-07171-2
2	83	12.6	191	5	PCT-US95-07171-3
3	83	12.6	892	1	US-07-977-434-12
4	64	9.7	892	1	US-08-456-819-12
5	64	9.7	892	1	PCT-US91-07035-14
6	64	9.7	3665	2	US-08-222-617A-1
7	64	9.8	3712	2	US-08-222-617A-4
8	64	9.8	3712	2	US-08-222-617A-25
9	64	9.7	120	1	US-08-347-462A-2
10	63.5	9.7	120	2	US-08-798-143-2
11	63.5	9.7	120	5	PCT-US95-15484-2
12	61	9.3	246	2	US-08-720-258-2

13	61	9.3	412	1	US-09-547-540A-8	Sequence 8, Appl
14	61	9.3	412	1	US-09-547-540A-9	Sequence 9, Appl
15	59	9.1	114	1	US-06-129-704-2	Sequence 2, Appl
16	59	9.1	114	1	US-06-172-604-2	Sequence 2, Appl
17	59	9.1	114	1	US-06-185-117-2	Sequence 2, Appl
18	59	9.1	114	1	US-06-177-537-2	Sequence 2, Appl
19	59	9.0	114	1	US-08-579-493A-15	Sequence 163, App
20	59	9.0	231	1	US-08-102-757-11	Sequence 11, Appl
21	59	9.0	1075	2	US-06-393-428-19	Sequence 19, Appl
22	59	9.0	1744	1	US-06-325-547-3	Sequence 1, Appl
23	58.5	8.9	355	2	US-06-683-743-4	Sequence 1, Appl
24	58.5	8.9	1211	1	US-09-134-001C-1820	Sequence 1820, Ap
25	58.5	8.9	1147	2	US-06-540-106-15	Sequence 19, Appl
26	58.5	8.9	1147	2	US-06-695-055-19	Sequence 19, Appl
27	58.5	8.9	1147	1	US-06-944-658-19	Sequence 19, Appl
28	58.5	8.9	1147	4	US-09-248-140-5	Sequence 5, Appl
29	58.5	8.9	1147	4	US-06-918-658-19	Sequence 15, Appl
30	58.5	8.9	1147	3	PCT-US95-13233-19	Sequence 15, Appl
31	58.5	8.9	2958	1	US-06-604-344C-2	Sequence 2, Appl
32	58.5	8.9	2958	1	US-09-542-249A-2	Sequence 2, Appl
33	57.5	8.6	429	4	US-09-312-658-2	Sequence 2, Appl
34	57	8.7	888	1	US-08-687-379-10	Sequence 10, Appl
35	57	8.7	888	1	US-08-687-379-12	Sequence 12, Appl
36	57	8.7	888	1	US-08-257-029-2	Sequence 2, Appl
37	57	8.7	888	4	US-06-257-029-4	Sequence 4, Appl
38	57	8.7	888	4	US-06-656-063-2	Sequence 4, Appl
39	57	8.7	888	4	US-08-895-063-4	Sequence 4, Appl
40	57	8.7	1444	2	US-08-540-106-10	Sequence 10, Appl
41	57	8.7	1444	2	US-08-656-055-10	Sequence 10, Appl
42	57	8.7	1434	4	US-08-954-668-10	Sequence 10, Appl
43	57	8.7	1434	4	US-06-918-658-10	Sequence 10, Appl
44	57	8.7	1434	2	PCT-US95-13233-19	Sequence 10, Appl
45	56.5	8.7	135	1	US-06-137-117D-100	Sequence 100, App

#### ALIGNMENTS

RESULT 1  
Sequence 2, Appl: action PC/US9507171  
GENERAL INFORMATION:  
APPLICANT: MCJNEY, J.  
TITLE OF INVENTION: Human Chemokine Beta-12  
NUMBER OF SEQUENCES: 7  
CORRESPONDENT ADDRESS:  
ADDRESSEE: CARILIA, BYRNIE, BAIN, GIFFILLIAN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
CREDITS: 2M 45/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PC/US95/07171  
FILING DATE: 6 JUN 95  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: none  
FILING DATE: none  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY W.  
REGISTRATION NUMBER: 350134  
REFERENCE/ATTORNEY NUMBER: 325000-391  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 136 AMINO ACIDS  
 TYPE: AMINO ACID  
 STRANDEDNESS:  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: PROTEIN  
 PCT-US95-07171-2

Alignment Scores:  
 Pred. No.: 0.00209 Length: 136  
 Score: 83.00 Matches: 24  
 Percent Similarity: 42.86% Conserved: 15  
 Best Local Similarity: 26.37% Mismatches: 32  
 Query Match: 12.63% Indels: 20  
 Gaps: 3

US-09-762-027-4 (1-388) x PCT-US95-07171-2 (1-136)

QY 136 CCGAGCTCGAGTGGATTTATGTTGTTGAGTCCGAGAGTATTAAGTAAAGAACTT 195  
 DB 5 ATGLeuSerLeuValIyrArgLysAlaLeuAlaGluArgMetLysValSerGluAlaIleu 24  
 QY 196 TCCATTGCCATCTG----- 210  
 DB 25 SerLeuValLeuIleLeuIleIleThrSerAlaSerArgSerGluProLysValPro 44  
 QY 211 -----AGAACTGGCAGCAGCAGCTGCTACAGCTAC---AAAGAACTTACAAI 261  
 DB 45 GluTrpValAsnThrProSerThrCysLeuLysTyrTyrGluLysValLeuProArg 64  
 QY 262 CATCTTCTTGAGGCA-----AGAACTGAGAGTCTACATAGAGCTTGCTTCATA 315  
 DB 65 ArgLeuValValIyrArgLysAlaLeuAlaCysHisLeuProAlaIleLeuPheVal 64  
 QY 316 ACTCTCAGAGCTGACAGATTCCTATATAATCCT 348  
 DB 85 ThrLysArgAsnArgGluValLysThrAsnPro 95

RESULT 2

PCT-US95-07171-3  
 Sequence 3 Application PCT/US9507171

GENERAL INFORMATION:

APPLICANT: MOONEY, J.

TITLE OF INVENTION: Human Chemokine Beta-12

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESS: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07171

FILING DATE: 6 JUN 95

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: None

FILING DATE: None

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-391

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 136 AMINO ACIDS  
 TYPE: AMINO ACID  
 STRANDEDNESS:  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: PROTEIN  
 PCT-US95-07171-2

Alignment Scores:  
 Pred. No.: 0.00209 Length: 136  
 Score: 83.00 Matches: 24  
 Percent Similarity: 42.86% Conserved: 15  
 Best Local Similarity: 26.37% Mismatches: 32  
 Query Match: 12.63% Indels: 20  
 Gaps: 3

US-09-762-027-4 (1-388) x PCT-US95-07171-3 (1-136)

QY 136 CCGAGCTCGAGTGGATTTATGTTGTTGAGTCCGAGAGTATTAAGTAAAGAACTT 195  
 DB 5 ATGLeuSerLeuValIyrArgLysAlaLeuAlaGluArgMetLysValSerGluAlaIleu 24  
 QY 196 TCCATTGCCATCTG----- 210  
 DB 25 SerLeuValLeuIleLeuIleIleThrSerAlaSerArgSerGluProLysValPro 44  
 QY 211 -----AGAACTGGCAGCAGCAGCTGCTACAGCTAC---AAAGAACTTACAAI 261  
 DB 45 GluTrpValAsnThrProSerThrCysLeuLysTyrTyrGluLysValLeuProArg 64  
 QY 262 CATCTTCTTGAGGCA-----AGAACTGAGAGTCTACATAGAGCTTGCTTCATA 315  
 DB 65 ArgLeuValValIyrArgLysAlaLeuAlaCysHisLeuProAlaIleLeuPheVal 64  
 QY 316 ACTCTCAGAGCTGACAGATTCCTATATAATCCT 348  
 DB 85 ThrLysArgAsnArgGluValLysThrAsnPro 95

RESULT 3

US-07-977-471  
 Sequence 3 Application US/07977134

GENERAL INFORMATION:

APPLICANT: Bellad, David H.

TITLE OF INVENTION: 5' to 3' LYMOULASE MUTATIONS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bellad, David H.

STREET: 340 Kingsland Street

CITY: New Jersey

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07977134

FILING DATE: 05/07/97

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 590,190

FILING DATE: 20-SEP-1990

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-391

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

```

      FILING DATE: 15-MAY-1990
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 143,441
      FILING DATE: 12-JAN-1988
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 063,509
      FILING DATE: 17-JUN-1987
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 899,241
      FILING DATE: 22-AUG-1986
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 746,121
      FILING DATE: 15-AUG-1991
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: WO PCT/US90/07641
      FILING DATE: 21-DEC-1990
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 565,471
      FILING DATE: 20-SEP-1990
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 455,611
      FILING DATE: 22-DEC-1985
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 609,157
      FILING DATE: 02-NOV-1990
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 557,517
      FILING DATE: 24-JUL-1990
      ATTORNEY/AGENT INFORMATION:
      NAME: Luann Cseri
      REGISTRATION NUMBER: 31,822
      REFERENCE/DOCKET NUMBER: Case No. 5466591 87535
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (510) 814-2572
      INFORMATION FOR SEQ ID NO: 12:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 892 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-07-577-434-12

```









Page 7

```

1  TITLE INVENTION: 11 EXPRESSED GENINES, THEIR
2  TITLE INVENTION: 11 PE DOCTIN AND USES
3  NUMBER: 1 SEQUENCES: 12
4  ADDRESS: 11111e Pharmaceut 415, Inc.
5  STREET: 1111 Porter Drive
6  CITY: 9410 Alib
7  STATE: VA
8  COUNTRY: U.S.
9  ZIP: 2104
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Diskette
12 OPERATING SYSTEM: DOS
13 SOFTWARE: FASTNU Version 1.5
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/750,113
16 FILING DATE: 1-FEB-1997
17 CLASSIFICATION: 536
18 PRICE AFFILIATION DATA:
19 APPLICATION NUMBER: 08/317,192
20 FILING DATE: 25-NOV-1994
21 APPLICATION NUMBER: 08/303,211
22 FILING DATE: 17-SEP-1994
23 APPLICATION NUMBER: 08/320,011
24 FILING DATE: 7-OCT-1994
25 AFOREMENTIONED INFORMATION:
26 NAME: Anthony, Barbara J
27 REGISTRATION NUMBER: 33,951
28 REFERENCE/BOOK NUMBER: PF-0024
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 415-955-0555
31 TELEFAX: 415-952-0195
32 INFORMATION FOR SEQ ID NO: 2:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 120 amino acids
35 TYPE: amino acid
36 STRANDNESS: single
37 TOPOLGY: linear
38 M LECOE TYPE: peptide
39 IMMEDIATE SOURCE:
40 LIBRARY: LIVER
41 COUNTRY: 27825
42 US-08-750-113-
43
44 Alignment Scores:
45 SEQ ID NO: 1
46 Score: 1.23
47 Percent Similarity: 63.50
48 Percent Identical: 47.924
49 Best Local Similarity: 33.334
50 Query Match: 5.674
51
52 Seqs: 2
53
54 05-08-752-007-1 (1-3ec) x 05-08-750-113-2 (1-120)
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
9
```

Fri Jan 3 14:26:59 2003

us-09-762-027-4.n2p.raii

Page 8

```

1  TITLE OF INVENTION:  US
2  NUMBER OF SEQUENCES:  6
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
5  STREET:  3174 PORTER DRIVE
6  CITY:  PALO ALTO
7  STATE:  CA
8  COUNTRY:  USA
9  ZIP:  94304
10
11  COMPUTER READABLE FORM:
12  MEDIUM TYPE:  Floppy disk
13  COMPUTER:  IBM PC compatible
14  OPERATING SYSTEM:  PC-DOS/MS-DOS
15  SOFTWARE:  Patent In Release #1.0, Version #1.30
16  CURRENT APPLICATION DATA:
17  APPLICATION NUMBER:  PCT/US95/15484
18  FILING DATE:  29-NOV-1995
19
20  CLASSIFICATION:
21  PRIOR APPLICATION DATA:
22  APPLICATION NUMBER:  US 08/347,452
23  FILING DATE:  29-NOV-1994
24
25  CLASSIFICATION:
26  ATTORNEY/AGENT INFORMATION:
27  NAME:  LUTHER, BARBARA J
28  REGISTRATION NUMBER:  33954
29  REFERENCE/DOCKET NUMBER:  PF-0024 PCT
30  TELECOMMUNICATION INFORMATION:
31  TELEPHONE:  415-855-0555
32  TELEFAX:  415-852-0195
33
34  INFORMATION FOR SEQ ID NO:  2:
35  SEQUENCE CHARACTERISTICS:
36  LENGTH:  120 amino acids
37  TYPE:  amino acid
38  STRANDEDNESS:  single
39  TOPOLOGY:  unknown
40  MOLECULE TYPE:  peptide
41  IMMEDIATE SOURCE:
42  LIBRARY:  liver
43
44  CLONE:  87825
45
46  PCT-US95-15484-2

```

```

Pred. No.: 1.23 Length: 120
Score: 63.50 Matches: 16
Percent Similarity: 47.92% Conservativeness: 7
Best Local Similarity: 33.33% Mismatches: 22
Query Match: 9.67% Indels: 3
DB: 5 Caps: 2

US-05-762-027-4 (1-386) x PCT-US95-15484-2 (1-120)

Cy 214 AACTGGCAGCAGGCAGCTGCTCCTACAGGCTAC---AAGGAAATTAGAAATCTTTT
      |||
      AsnTrpProSerThrCysCysLeuStyIleValysValLeuProArgGlyLeuVal
Db 32 AsnTrpProSerThrCysCysLeuStyIleValysValLeuProArgGlyLeuVal
Cy 271 GAGGGA-----AACCACTGAGAGCGCTACATAGATGCTGCTGTTTCAACTCTGCA
      |||
      ValGlyTyrArgLysAlaLeuIleuLysCysIleuProAlaIleIlePheValIleLysArg
Db 52 ValGlyTyrArgLysAlaLeuIleuLysCysIleuProAlaIleIlePheValIleLysArg
Cy 325 AGCTGCAGATTCGTGTATAAATCTT 348
      |||
      AsnArgGluValCysThrAsnPro 79
Db 72 AsnArgGluValCysThrAsnPro 79

RESULT 12
: US-05-720-258-2
: Sequence 2, Application US/08720258
: Patent No. 5871740
: GENERAL INFORMATION:
: APPLICANT: Smith, Craig A.
: TITLE OF INVENTION: Chemokine Inhibitor
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation

```

1 NAME: Steven E. Gonsky  
 2 Address: Seattle  
 3 City: Washington  
 4 Country: US  
 5 ZIP CODE: 98101  
 6 REPORT REQUESTED FROM:  
 7 MAILING TYPE: floppy disk  
 8 MEDIA: Apple Power Macintosh  
 9 FILING SYSTEM: Apple System 7.5.3  
 10 SUBJECT: Microsoft Word, Version 6.0.1  
 11 JOURNAL APPLICATION DATA:  
 12 APPLICATION NUMBER: 05/06/720,426  
 13 FILING DATE: 29-SEP-1995  
 14 CLASSIFICATION: 445  
 15 FILE APPLICATION DATA:  
 16 APPLICATION NUMBER: US-06/575,715  
 17 FILING DATE: 20-DEC-1995  
 18 PATENT APPLICATION DATA:  
 19 APPLICATION NUMBER: US-06/537,324  
 20 FILING DATE: 29-SEP-1995  
 21 AIR REPLY/AGENT INFORMATION:  
 22 NAME: Anderson, Katherine A.  
 23 REGISTRATION NUMBER: 172,172  
 24 REFERENCE/DOSKEY NUMBER: 222,008  
 25 TELEPHONIC/INFORMATION:  
 26 TELEPHONE: (206) 567,7450  
 27 TELEFAX: (206) 233-0044  
 28 TELEX: 756622  
 29 INFORMATION FOR SEQ ID NO.:  
 30 SEQUENCE CHARACTERISTICS:  
 31 LENGTH: 419 amino acids  
 32 TYPE: amino acid  
 33 ORIGIN: linear  
 34 MAGNITUDE TYPE: protein  
 35 OS-067537-003-2

[illegible]



[illegible]

```
-Search completed: January 2, 2003, 15:13:30
Job time : 16.5 secs
```

Fri Jan 3 14:26:59 2003

us-09-762-027-4.n2p.rapb

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

us nucleic - protein search, using frame\_plus\_n2p model

Run on: January 2, 2003, 15:08:03 : Search time 49 seconds  
(without alignment)

735,319 Million cell updates/sec

File: US-09-762-027-4  
Perfect score: 657  
Sequence: 1 agctgcataacaggaacac.....atccataaagatataacac 108

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 117078 seqs, 18951520 residues

Total number of hits satisfying chosen parameters: 41156

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_n2p\_model -DEV=xlh  
-Q=/cgn2\_1/uspro.spool/US09762027/runal\_02012003\_085614\_2682/epc\_qnery.fasta\_1.55  
-DB=Published.Applications\_AA -QEMF=FASTAN -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-ITER\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NCRN=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09762027\_1/cgn2\_1\_1\_2\_3/runal\_02012003\_085614\_2682  
-NCPU=6 -ICPU=3 -NO\_MLPHY -NO\_MMAP -LARGEOUTPUT -NEE\_SCORES=0 -MATT=1 -LARGEOUT3  
-DEV\_1TIMEOUT=120 -WARN\_1TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=0  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA:

- 1: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/prodata/2/pubppa/PC1\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep.\*
- 5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubppa/PCUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep.\*
- 9: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep.\*
- 11: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	91	13.9	241	10	US-09-925-300-1065 Sequence 1065, App1
2	77	11.7	181	10	US-09-867-950-930 Sequence 930, App1
3	61.5	9.4	372	10	US-09-768-894A-2 Sequence 2, App1
4	60	9.1	276	10	US-09-943-798-2 Sequence 2, App1

Sequence	Score	Match	Length	DB ID	Description
Sequence 1, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 10727, A	77	11.7	181	10	US-09-867-950-930
Sequence 20, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 6, App1	60	9.1	276	10	US-09-943-798-2
Sequence 23, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 43, App1	77	11.7	181	10	US-09-867-950-930
Sequence 41, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 10, App1	60	9.1	276	10	US-09-943-798-2
Sequence 25, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 137, App1	77	11.7	181	10	US-09-867-950-930
Sequence 3, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 5, App1	60	9.1	276	10	US-09-943-798-2
Sequence 360, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 1057, A	77	11.7	181	10	US-09-867-950-930
Sequence 1923, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 5, App1	60	9.1	276	10	US-09-943-798-2
Sequence 26, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 27, App1	77	11.7	181	10	US-09-867-950-930
Sequence 17, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 2, App1	60	9.1	276	10	US-09-943-798-2
Sequence 376, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 12291, A	77	11.7	181	10	US-09-867-950-930
Sequence 4, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 3, App1	60	9.1	276	10	US-09-943-798-2
Sequence 13, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 2, App1	77	11.7	181	10	US-09-867-950-930
Sequence 8, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 6, App1	60	9.1	276	10	US-09-943-798-2
Sequence 30, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 1, App1	77	11.7	181	10	US-09-867-950-930
Sequence 2, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 9, App1	60	9.1	276	10	US-09-943-798-2
Sequence 24, App1	91	13.9	241	10	US-09-925-300-1065
Sequence 22, App1	77	11.7	181	10	US-09-867-950-930
Sequence 242, App1	61.5	9.4	372	10	US-09-768-894A-2
Sequence 105, App1	60	9.1	276	10	US-09-943-798-2

## ALIGNMENTS

Result 1  
US-09-925-300-1065  
Sequence 1, App1  
US-09-925-300-1065  
GENERAL INFORMATION:  
APPLICANT: GENE ROSEN  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
CURRENT INVENTION: Nucleic Acids, Proteins and Antibodies  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1997-03-12  
NUMBER OF SEQ. ID NOS: 1690  
SEQUENCE ID NOS: 1690  
LENGTH: 241  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-300-1065  
Alignment Scores:  
Pred. No.: 91  
Score: 13.9  
Percent Similarity: 91.00  
Best Local Similarity: 91.14  
Length: 241  
Matches: 20  
Conservative: 7  
Mismatches: 7



```

Query Match: 13.65%   Models: 1
DB: 10               Gaps: 0

US-09-762-027-4 (1-388) x US-09-925-300-1065 (-211)

QY 282 GCTGAAGTCTCATAGATGCTGTGCTTCAATATCTCAGAGCTGAGATTCTTAT 341
DB 9 SerGialSerHisArgMetIleLeuLeuAsnSerHisLysLeuLeuAlaLeuTyr 26

QY 342 AAATCCTGAAACAGACATCCCTGATCATATCAATATATAT 384
DB 29 LysSerLeuAlaArgSerIle-ProLysLeuLysValTyr 42

RESULT 2
US-09-867-550-930
Sequence 950, Application US/09867550
Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad.
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1e] Polynucleotides from Atherogenetic cells and
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 930
LENGTH: 181
TYPE: PRI
ORGANISM: Homo sapiens
US-09-867-550-930

Alignment Scores:
Pred. No.: 0.0156   Length: 181
Score: 77.60       Matches: 16
Percent Similarity: 79.31%   Conservative: 5
Best Local Similarity: 62.07%   Mismatches: 5
Query Match: 11.72%   Indels: 1
DB: 10             Gaps: 0

US-09-762-027-4 (1-388) x US-09-867-550-930 (1-181)

QY 300 ATGCTTGCTTCATATACCTGACAGAGCTGCAGATTCTATATATCTTACGAAATAT 359
DB 1 MetIleLeuLeuAsnSerHisLysLysLeuAlaLeuTyrLysSerIleAlaAlaTser 2

QY 360 ATCCCTGATCATATCAATATATAT 384
DB 21 Ile-ProLysLeuLysValTyr 28

RESULT 3
US-09-768-894A-2
Sequence 2, Application US/09768894A
Patent No. US20020115102A1
GENERAL INFORMATION:
APPLICANT: Diane Joan Couzens
APPLICANT: Steven Michael Foord
APPLICANT: Diane Michele Ignar
APPLICANT: Filippo Volpe
TITLE OF INVENTION: NOVEL PROTEIN
FILE REFERENCE: PG37700S
CURRENT APPLICATION NUMBER: US/09/768,894A
CURRENT FILING DATE: 2002-02-24
PRIOR APPLICATION NUMBER: 09/768,894
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1

```

```

SEQ ID NO 930
LENGTH: 181
TYPE: PRI
ORGANISM: Homo sapiens
US-09-768-894A-2

Alignment Scores:
Pred. No.: 0.0156   Length: 181
Score: 77.60       Matches: 16
Percent Similarity: 79.31%   Conservative: 5
Best Local Similarity: 62.07%   Mismatches: 5
Query Match: 11.72%   Indels: 1
DB: 10             Gaps: 0

US-09-762-027-4 (1-388) x US-09-768-894A-2 (1-372)

QY 300 ATGCTTGCTTCATATACCTGACAGAGCTGCAGATTCTATATATCTTACGAAATAT 359
DB 1 MetIleLeuLeuAsnSerHisLysLysLeuAlaLeuTyrLysSerIleAlaAlaTser 2

QY 360 ATCCCTGATCATATCAATATATAT 384
DB 21 Ile-ProLysLeuLysValTyr 28

RESULT 4
US-09-943-798-2
Sequence 2, Application US/09943798
Patent No. US20020071798
GENERAL INFORMATION:
APPLICANT: Glaxo Group Limited
TITLE OF INVENTION: No. US20020071798
FILE REFERENCE: 031-123
CURRENT APPLICATION NUMBER: US/09/943,798
CURRENT FILING DATE: 2001-02-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 276
TYPE: PRI
ORGANISM: Homo sapiens
US-09-943-798-2

Alignment Scores:
Pred. No.: 0.0156   Length: 276
Score: 77.60       Matches: 26
Percent Similarity: 79.31%   Conservative: 7
Best Local Similarity: 62.07%   Mismatches: 17
Query Match: 11.72%   Indels: 6
DB: 10             Gaps: 2

US-09-762-027-4 (1-388) x US-09-943-798-2 (1-276)

QY 140 GATTTTATATGCTTCTATATATATATATATATATATATATATATATATATATAT 192
DB 221 GATTTTATATGCTTCTATATATATATATATATATATATATATATATATATATAT 238

QY 193 GATTTTATATGCTTCTATATATATATATATATATATATATATATATATATATAT 249
DB 439 GATTTTATATGCTTCTATATATATATATATATATATATATATATATATATATAT 258

QY 250 GATTTTATATGCTTCTATATATATATATATATATATATATATATATATATATAT 258

```

Db 259 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 270

RESULT 5  
US-09-943-798-4  
Sequence 4, Application US/09943798  
Patent No. US20020065215A1

GENERAL INFORMATION:

APPLICANT: Glaxo Group Limited

TITLE OF INVENTION: Polypeptide

FILE REFERENCE: Q61021

CURRENT APPLICATION NUMBER: US/09/943,798

CURRENT FILING DATE: 2001-08-31

NUMBER OF SEQ. ID NOS: 4

SOFTWARE: FastSeq for Windows Version 3.0

SEQ. ID NO. 4

LENGTH: 337

TYPE: PRI

ORGANISM: Homo sapiens

US-09-943-798-4

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-762-027-4 (1-388) x US-09-943-798-4 (1-337)

QY 148 GAATTTATTGCTTTCAGTCCGACAGTATTTACCTACAGCAAC-----

Db 282 GlnAlaIyrIleValSer-----ArgProLeuAlaIleValSerThrPheGlyAsnLeu 299

QY 153 ---CTTTCATTCGCTACTCTGAGAACTGGCAGCAGGAGCTGTCCCTACAGCTACAAA 243

Db 300 LeuLeuIyrValIleValSerAspAsnIleGlnAlaValCysSerThrValAspGly 319

QY 250 GAACCTCAGATCATCTCTCTGAGGAGGAAGACGCT 285

Db 320 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 331

RESULT 6

US-09-864-761-48727

Sequence 48727, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: A60000-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

Db 150 LysValSerGlyAsnLeuGluGlnAlaLysLysIle 160

RESULT 7

US-09-771-955-4

Sequence 20, Appl. No. US/09771956

Patent No. US-09-771,956

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: A60000-X-1

CURRENT APPLICATION NUMBER: US/09/771,956

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30









US-09-762-027-4 (1-388) x US-10-066-500-137 (1-325)

QY 154 TATGTGTTACGTCGCTAGAGATTAGCTAACCAAAACCTTCGATTGCAATGAGAA 213  
||||| ||||| ||| :||| :  
Db 187 TTYCYValGlnValArgGlyPheLeuProAspArgAsnLysAla-----Gly 242

QY 214 AACGGCAGCAGCAGCTGCTACAG---GTCTACAAAGAACCTCAGATCATCTTCT 27  
||| :||| ||| :||| :  
Db 203 GlnTrpSerGlnProValCysGlnGlnThrHisAspGlnThrValProSerTrpMet 322

QY 271 GAGCGAAGAGCTG-----AAGTCTACATAAGAAAGCTTCTCTTCAATCTCAGA 24  
||| :||| :  
Db 223 ValAlaValAlaLeuMetAlaSerValPheMetValCysLeuAlaLeuGlyCysPhe 242

QY 325 AACGGCAGATTCTGTAFA 342  
||| :||| :  
Db 243 SerLeuLeuTrpCysVal 248

## RESULT 15

US-09-762-027-4

: Sequence 3, Application US/09870574

: Patent No. US20020102723A1

: GENERAL INFORMATION:

: APPLICANT: Gurney, Austin L.

: APPLICANT: Aggarwal, Sudeepa

: APPLICANT: Xie, Ming-Hong

: APPLICANT: Maruoka, Ellen M.

: APPLICANT: Foster, Jessica S.

: APPLICANT: Goddard, Audrey

: APPLICANT: Wood, William I.

: TITLE OF INVENTION: INTERLEUKIN-22 POLYPEPTIDES, NUCLEIC ACIDS AND ANTIBODIES

: FILE OF INVENTION: THE SAME AND METHODS FOR THE TREATMENT OF PANCREATIC DISORDERS

: FILE REFERENCE: P2806-1(US)

: CURRENT APPLICATION NUMBER: US/09/870,574

: CURRENT FILING DATE: 2001-05-30

: PRIOR APPLICATION NUMBER: US 60/169,495

: PRIOR FILING DATE: 1999-12-07 PCT/US00/14042

: PRIOR APPLICATION NUMBER: PCT/US00/23328

: PRIOR FILING DATE: 2000-08-24

: NUMBER OF SEQ ID NOS: 7

: SEQ ID NO: 3

: LENGTH: 325

: TYPE: PRT

: ORGANISM: Homo Sapien

: US-09-762-027-4

## Alignment Scores:

Prod. No.:	8.4	Length:	325
Score:	58.50	Matches:	19
Percent Similarity:	43.94%	Conservative:	10
Best Local Similarity:	28.79%	Mismatches:	30
Query Match:	8.90%	Indels:	7
DB:	10	Gaps:	3

US-09-762-027-4 (1-388) x US-09-870-574-3 (1-325)

QY 154 TATGTGTTACGTCGCTAGAGATTAGCTAACCAAAACCTTCGATTGCAATGAGAA 213  
||||| ||||| ||| :||| :  
Db 187 TTYCYValGlnValArgGlyPheLeuProAspArgAsnLysAla-----Gly 242

QY 214 AACGGCAGCAGCAGCTGCTACAG---GTCTACAAAGAACCTCAGATCATCTTCT 27  
||| :||| ||| :||| :  
Db 203 GlnTrpSerGlnProValCysGlnGlnThrHisAspGlnThrValProSerTrpMet 322

QY 271 GAGCGAAGAGCTG-----AAGTCTACATAAGAAAGCTTCTCTTCAATCTCAGA 24  
||| :||| :  
Db 223 ValAlaValAlaLeuMetAlaSerValPheMetValCysLeuAlaLeuGlyCysPhe 242

QY 325 AACGGCAGATTCTGTAFA 342  
||| :||| :  
Db 243 SerLeuLeuTrpCysVal 248

Db 243 SerLeuLeuTrpCysVal 248

Search completed January 2, 2003, 15:12:54  
Job time: 21 secs